Access DB# 73646

SEARCH REQUEST FORM

Scientific and Technical Information Center

Requester's Full Name: _____ Examiner # :_____ Date: _____
Art Unit: _____ Phone Number 30_____ Serial Number: _____
Mail Box and Bldg/Room Location: _____ Results Format Preferred (circle): PAPER DISK E-MAIL

If more than one search is submitted, please prioritize searches in order of need.

****

Please provide a detailed statement of the search topic, and describe as specifically as possible the subject matter to be searched. Include the elected species or structures, keywords, synonyms, acronyms, and registry numbers, and combine with the concept or utility of the invention. Define any terms that may have a special meaning. Give examples or relevant citations, authors, etc, if known. Please attach a copy of the cover sheet, pertinent claims, and abstract.

Title of Invention: _____

Inventors (please provide full names): _____

_____

Earliest Priority Filing Date: _____

***For Sequence Searches Only** Please include all pertinent information (parent, child, divisional, or issued patent numbers) along with the appropriate serial number.*

Jan Delaval
Reference Librarian
Biotechnology & Chemical Library
CM1 1E07 – 703-308-4498
jan.delaval@uspto.gov

****

STAFF USE ONLY	Type of Search	Vendors and cost where applicable
Searcher: _____	NA Sequence (#)_____	STN _____
Searcher Phone #: 4495	AA Sequence (#) 16	Dialog
Searcher Location:	Structure (#)	Questel/Orbit
Date Searcher Picked Up: 8/26/02	Bibliographic	Dr.Link
Date Completed: 8/26/02	Litigation	Lexis/Nexis
Searcher Prep & Review Time:	Fulltext	Sequence Systems ✓
Clerical Prep Time: 20	Patent Family	WWW/Internet
Online Time: 15	Other	Other (specify)

PTO-1590 (8-01)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

GM protein - protein search, using sw model

Run on: August 26, 2002, 06:10:50 ; Search time 80.5 Seconds
(without alignments)
15,518 Million cell updates/sec

Title: US-09-744-282-6
Perfect score: 62
Sequence: 1 RSFTLASSETGVG 13

Scoring table: BLOSUM62
Gap p 13.0 , Gapext 0.5

Searched: 28148 seqs, 9618334 residues

Total number of hits satisfying chosen parameters: 28138

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 71:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	62	100.0	383	2	A49562	cartilage glycopro
2	51	82.3	383	2	S51327	heparin binding gl
3	43	69.4	405	2	S61561	breast regressing
4	41	66.1	830	2	T09578	1,4-alpha-glucan b
5	40	64.5	185	2	S50094	endodeoxyribonucle
6	40	64.5	461	2	G90791	hypothetical prote
7	40	64.5	467	2	D85652	hypothetical prote
8	39	62.9	587	2	G70502	probable recN prot
9	38	61.3	261	2	F84325	hypothetical prote
10	38	61.3	286	2	H98250	lysk type transcri
11	38	61.3	286	2	AC4045	transcription regu
12	38	61.3	441	2	T41090	probable ATP-bindi
13	37	59.7	242	2	AE0626	conserved hypothet
14	37	59.7	298	2	D95967	probable transcrip
15	37	59.7	459	2	F84895	metal-tetracycline
16	36	58.1	251	2	C96846	conserved hypothet
17	36	58.1	266	2	C70798	hypothetical prote
18	36	58.1	424	2	G81450	probable aminotran
19	36	58.1	664	2	AG1026	topoisomerase B [i
20	36	58.1	860	2	T49910	hypothetical prote
21	36	58.1	861	2	T49909	hypothetical prote
22	36	58.1	1555	2	T18688	hypothetical prote
23	36	58.1	1973	2	T18686	hypothetical prote
24	36	58.1	1973	2	G89608	protein B0272.5 [i
25	36	58.1	2305	2	T15571	hypothetical prote
26	36	58.1	2305	2	BB9608	protein C23F12.1 [
27	35	56.5	187	2	S69542	hypothetical prote
28	35	56.5	271	2	H97230	alpha/beta superta
29	35	56.5	299	2	E98151	transcription regu
30	35	56.5	299	2	AD3156	transcriptional reg
31	35	56.5	302	2	F95387	protein (imported
32	35	56.5	320	2	T24905	hypothetical prote
33	35	56.5	355	2	H97715	alanine racemase (
34	35	56.5	383	2	B87044	MRP-family ATP bin
35	35	56.5	414	2	T43387	hypothetical prote
36	35	56.5	433	2	GB4612	probable GIP cyclo
37	35	56.5	445	2	T45199	probable mrp prote
38	35	56.5	445	2	B40590	flagellar distal c
39	35	56.5	573	2	B71424	probable DNA repai
40	35	56.5	596	2	AC1418	two component sens
41	35	56.5	664	2	G98422	hypothetical prote
42	35	56.5	664	2	AG2960	glycosidase A [ri
43	35	56.5	686	2	G64618	cadmium transporti
44	35	56.5	686	2	E71895	probable heavy met
45	34	54.8	194	2	A84081	GIP binding protei

ALIGNMENTS

RESULT 1
A49562
cartilage glycoprotein gp39 precursor - human
N;Alternate names: 39K synovial protein
C;Species: Homo sapiens (man)
C;Date: 23-Mar-1995 #sequence_revision 23-Mar-1995 #text_change 22-Jun-1999
C;Accession: A49562; S10677; A33162
R;Hakala, B.E.; White, C.; Recklies, A.D.
J. Biol. Chem. 268, 25803-25810, 1993
A;Title: Human cartilage gp-39, a major secretory product of articular chondrocytes a
A;Reference number: A49562; MUID:94064658
A;Accession: A49562
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-383 <HAK>
A;Cross-references: GB:M80927; NID:g348911; PIDN:AAA16074.1; PID:g348912
R;Nyirkos, P.; Golds, E.E.
Biochem. J. 269, 265-268, 1990
A;Title: Human synovial cells secrete a 39 kDa protein similar to a bovine mammary pr
A;Reference number: S10677; MUID:90328983
A;Accession: S10677
A;Molecule type: protein
A;Residues: 22-40,'X',42-45 <NY2>
C;Superfamily: Streptomyces chitinase chi40
C;Keywords: cartilage; extracellular protein; glycoprotein
F;1-21/Domain: signal sequence #status predicted <SIG>
F;22-383/Product: cartilage glycoprotein gp39 #status predicted <MAT>

Query Match 100.0%; Score 62; DB 2; Length 383;
Best Local Similarity 100.0%; Prod. No. 0.00046;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RSFTLASSETGVG 13
||||||||||||| |
Db 264 RSFTLASSETGVG 276

RESULT 2
S51327
heparin binding glycoprotein 38K - pig
C;Species: Sus scrofa domestica (domestic pig)
C;Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 22-Jun-1999
C;Accession: S51327
R;Shackelton, L.M.; Mann, D.M.; Millis, A.J.T.
submitted to the EMBL Data Library, January 1995
A;Description: Identification of a 38kDa heparin binding glycoprotein (gp38k) in diff
A;Reference number: S51327
A;Accession: S51327
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-383 <SHA>

A;Cross-references: EMBL:Z47804; NID:g1345957; PIDN:CAA87943.1; PID:g1345958
C;Superfamily: Streptomyces chitinase chi40

Query Match 82.4%; Score 51; DB 2; Length 88;
Best Local Similarity 84.6%; Pred. No. 0.052;
Matches 11; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RSFTLASSETGVC 13
| | | | | | |
Db 263 RSFTLASSKTDVC 275

RESULT 3
S61551
breast-retrossing protein brp39 precursor - mouse
C;Species: Mus musculus (house mouse)
C;Date: 16-Mar-1995 #sequence_revision 25-Apr-1997 #text_change 20-Jun-2000
C;Accession: S61551; S62550; 148271
R;Morrison, B.W.; Leder, P.
Oncogene 9, 3417-3426, 1994
A;Title: neu and ras initiate murine mammary tumors that share genetic markers generally
A;Reference number: 148271; MUID:95006747
A;Accession: S61551
A;Status: nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1-405 <MOR>
A;Cross-references: EMBL:X93035; NID:g1085005; PIDN:CAA63555.1; PID:g1085006
R;Morrison, B.W.
submitted to the EMBL Data Library, November 1995
A;Reference number: S61551
A;Accession: S61551
A;Molecule type: mRNA
A;Residues: 1-247,'I',247-330,'H',332-350,'MVWALHLIEDFQSIC',368-EEF-FI-INAIKDALA' <MOR2>
A;Cross-references: EMBL:X92035; NID:g1085004; PIDN:CAA63554.1; PID:g1085005
A;Note: the differences at the carboxyl end are due to a frameshift error
C;Genetics:
A;Gene: brp39
C;Superfamily: Streptomyces chitinase chi40
F;1-21/Domain: signal sequence #status predicted <SIG>
F;22-405/Product: breast-retrossing protein brp39 #status predicted <MAT>

Query Match 69.4%; Score 43; DB 2; Length 88;
Best Local Similarity 69.2%; Pred. No. 2.1;
Matches 9; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 RSFTLASSETGVC 13
| | | | | | |
Db 264 RSFTLASSENQTC 276

RESULT 4
I06578
1,4-alpha-glucan branching enzyme (EC 2.4.1.18) chel precursor - wheat
N;Alternate names: starch branching enzyme I
C;Species: Triticum aestivum (common wheat)
C;Date: 25-Apr-1999 #sequence_revision 25-Apr-1999 #text_change 20-Jun-2000
C;Accession: I06578
R;Chibbar, R.N.
submitted to the EMBL Data Library, April 1997
A;Reference number: Z15772
A;Accession: I06578
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-836 <HU>
A;Cross-references: EMBL:Y12320; PIDN:CAA72987.1
A;Experimental source: cv. Fielder; kernels at 12 days post anthesis
C;Genetics:
A;Gene: chel
C;Function:
A;Description: catalyzes the transglycosylation of a terminal segment of a 1,4-alpha-D-g
A;Pathway: glycogen/starch biosynthesis
C;Superfamily: 1,4-alpha-glucan branching enzyme
C;Keywords: amyloplast; glycogen/starch biosynthesis; glycosyltransferase; transferase
F;1-61/Domain: transit peptide (amyloplast) #status predicted <TNP>
F;62-830/Product: 1,4-alpha-glucan branching enzyme chel #status predicted <MAT>

Query Match 66.1%; Score 41; DB 2; Length 836;
Best Local Similarity 83.3%; Pred. No. 11;
Matches 7; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

Qy 1 RSFTLASSETGVC 13
| | | | | | |
Db 66 RDYTMATAEDSVG 78

RESULT 5
S50094
cytodeoxyribonuclease (EC 3.1.21.-) - phage SP8
C;Species: phage SP8
A;Note: host Bacillus subtilis
C;Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 20-Jun-2000
C;Accession: S50094
R;Goodrich-Blair, H.; Shub, D.A.
Nucleic Acids Res. 22, 3715-3721, 1994
A;Title: The DNA polymerase genes of several BMU bacteriophages have a unique structure
A;Reference number: S50094; MUID:95052417
A;Accession: S50094
A;Molecule type: DNA
A;Residues: 1-185 <Goo>
A;Cross-references: EMBL:U24414; NID:g434134; PIDN:AAA88284.1; PID:g434135
C;Superfamily: cytosine deoxyribonuclease
C;Keywords: hydrolase

Query Match 63.7%; Score 4...; DB 2; Length 3...;
Best Local Similarity 75.0%; Pred. No. ...4;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 RSFTLASSETGV 12
| | | | | | | | | | | | |
Db 144 RSITQASKEIGV 155

RESULT 6
G90791
hypothetical protein b3518 [imported] - Escherichia coli (strain K-12, substrain
C;Species: Escherichia coli
C;Date: 18-Jul-2001 #sequence_revision 20-Jul-2001 #text_change ...-Aug-...
C;Accession: G90791
R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C
Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and g
A;Reference number: A90629; MUID:21156241; PMID:11258796
A;Accession: G90791
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-461 <HAY>
A;Cross-references: GB:BA000007; PIDN:BAB34726.1; PID:g13361264; GSPDB:GN00174
A;Experimental source: strain O157:H7, substrain RIMD 0509952
C;Genetics:
A;Gene: ECs1403
C;Superfamily: Escherichia coli hypothetical protein b3518

Query Match 64.5%; Score 4...; DB 2; Length 4...;
Best Local Similarity 75.0%; Pred. No. 9.4;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 RSFTLASSETGV 12
| | | | | | | | | | | | |
Db 278 RSFALVSSESGV 289

RESULT 7
D85652
hypothetical protein [imported] - Escherichia coli (strain O157:H7, substrain EDL9
C;Species: Escherichia coli
C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
C;Accession: D85652; D85601
R;Perna, N.T.; Plunkett, G., 3rd; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
Nature 409, 529-533, 2001
A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A;Reference number: A85480; MUID:21074935; PMID:11206551
A;Accession: D85652
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-467 <PE>
A;Cross-references: GB:AE005174; NID:g12514439; PIDN:AAG55680.1; GSPDB:GN00145; UWGP:Z15
A;Experimental source: strain O157:H7, substrain EDL933
A;Accession: D85601
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-467 <SE2>
A;Cross-references: GB:AE005174; NID:g12513922; PIDN:AAG55271.1; GSPDB:GN00115; UWGP:Z11
A;Experimental source: strain O157:H7, substrain EDL933
C;Genetics:
A;Gene: Z1565; Z1126
C;Superfamily: Escherichia coli hypothetical protein b2625

Query Match 64.5%; Score 40; DB 2; Length 467;
Best Local Similarity 75.0%; Pred. No. 9.5;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 RSFTLASSETVG 12
||| : ||| ||
Db 184 RSFALVSSEQTV 195

RESULT 8
G70502
probable recN protein - Mycobacterium tuberculosis (strain H37Rv)
C;Species: Mycobacterium tuberculosis
C;Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 20-Jun-2001
C;Accession: G70502
R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S
; Eiglmeier, K.; Gas, S.; Barry, C.E. 3rd; Tekaia, F.; Badcock, K.; Basham, D.; Brown,
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A;Authors: Spares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A;Reference number: A70500; MUID:98295987
A;Accession: G70502
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-587 <SOL>
A;Cross-references: GB:Z95208; GB:AL123456; NID:g2621833; PIDN:CAB10553.1; GSPDB:GN00
A;Experimental source: strain H37Rv
C;Genetics:
A;Gene: recN
C;Superfamily: recN protein

Query Match 62.9%; Score 39; DB 2; Length 587;
Best Local Similarity 54.8%; Pred. No. 19;
Matches 7; Conservative 2; Mismatches 4; Indels 3; Gaps 0;

Qy 1 RSFTLASSETVG 13
|:|:| : | | |
Db 22 RGFIVLIGETIG 34

RESULT 9
F84325
hypothetical protein VNG1735C [imported] - Halobacterium sp. NRC-1
C;Species: Halobacterium sp. NRC-1
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
C;Accession: F84325
R;Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky
; Baliga, N.S.; Thorsson, V.; Sbrogna, J.; Swartzell, S.; Weir, D.; Hall, J.; Dahl,
Hood, L.; Hershberg, R.; Haynes, P.A.; Raymond, J.; Glusman, G.; Kennedy, K.; Schmid,
Hackett, M.; Bolanos, R.; Faith, J.; Ward, D.; Hoff, K.; Pan, M.; Walian, P.; Buican,
Kim, C.; Facciotti, M.T.; Burnette, D.; Kim, J.S.; Goo, Y.A.; Sharma, V.; Kan, Z.; Ab
Jung, K.H.; Alam, M.; Freitas, T.
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
A;Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.;
A;Title: Genome sequence of Halobacterium species NRC-1.
A;Reference number: A84168; MUID:20504483
A;Accession: F84325
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-261 <SIO>
A;Cross-references: GB:AE004437; NID:g10581197; PIDN:AAG19970.1; GSPDB:GN00138
C;Genetics:
A;Gene: VNG1735C

Query Match 61.3%; Score 38; DB 2; Length 261;
Best Local Similarity 63.6%; Pred. No. 12;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 RSFTLASSETG 11
|:|:| |:||
Db 35 RIFTLVDSDIG 45

RESULT 10
H98250
lysR-type transcription regulator [imported] - Agrobacterium tumefaciens (strain C58,
C;Species: Agrobacterium tumefaciens
C;Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 11-Jan-2002
C;Accession: H98250
R;Goodner, B., Hinkle, G., Gattung, S., Miller, N., Blanchard, M., Qurollo, B., Goldm
A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz,
Science 294, 2323-2328, 2001
A,Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium
A;Reference number: A97359; PMID:11743194
A;Accession: H98250
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-286 <KUR>
A,Cross references: GB:AE007870; PIDN:AAK89530.1; PID:g15594409; GSPDB:GN00170
C;Genetics:
A;Gene: AGR_L_1915
A;Map position: linear chromosome
C;Superfamily: hypothetical protein bl1875

Query Match 61.3%; Score 38; DB 2; Length 286;
Best Local Similarity 58.8%; Pred. No. 14;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 2 SFTLASSETGVG 13
|||||: |:|
Db 18 SFTLAAERLGIG 29

RESULT 11
AC3045
transcription regulator, lysR family Atu3889 [imported] - Agrobacterium tumefaciens (
C;Species: Agrobacterium tumefaciens
C;Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 01-Feb-2002
C;Accession: AC3045
R;Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo
erage, G.; Gillet, W.; Grant, C.; Guenthner, D.; Kutyavin, I.; Levy, R.; Li, M.; McC
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A;Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kam
ster, E.W.
A;Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.

A;Reference number: AB2577; PMID:11744193
A;Accession: AC0035
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-286 <KUR>
A;Cross-references: GB:AE000689; PIDN:AAL44697.1; PID:g17742526; GSPDB:GN00287
A;Experimental source: strain C58 (Dupont)
C;Genetics:
A;Gene: Atu3889
A;Map position: linear chromosome
C;Superfamily: hypothetical protein bl875

Query Match 61.3%; Score 38; DB 2; Length 286;
Best Local Similarity 58.3%; Pred. No. 14;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 2 SFTLASSETGVG 13
:|| : : |:|
Db 98 SFTAAEKDGIT 29

RESULT 12
I41090
probable ATP-binding protein - Serratia marcescens
C;Species: Serratia marcescens
C;Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 17-Mar-2000
C;Accession: I41090
R;Saigi, F.; Climent, N.; Pique, N.; Sanchez, C.; Merino, S.; Rubires, X.; Aguilar, A.;
J. Bacteriol. 181, 1883-1891, 1999
A;Title: Genetic analysis of the Serratia marcescens N28b O4 antigen gene cluster.
A;Reference number: S20974; MUID:99173913
A;Accession: I41090
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-441 <SAI>
A;Cross-references: EMBL:AF038816; NID:g2828669; PID:g2828672; PIDN:AAC00182.1
C;Genetics:
A;Gene: wzt
C;Superfamily: unassigned ATP binding cassette proteins; ATP binding cassette homology

Query Match 61.3%; Score 38; DB 2; Length 441;
Best Local Similarity 88.9%; Pred. No. 22;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 SFTLASSETGVG 13
:|||||||||||
Db 331 SFTLASSETGVG 344

RESULT 13
A69026
conserved hypothetical protein MTH1192 - Methanobacterium thermoautotrophicum (strain DH
C;Species: Methanobacterium thermoautotrophicum
C;Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 22-Oct-1999
C;Accession: A69026
R;Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.;
; Qiu, D.; Spadafora, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiwani, N.
ki, S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Nelson, J.; Reeve, J.N.
J. Bacteriol. 179, 7135-7155, 1997
A;Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: funct
A;Reference number: A69000; MUID:98037514
A;Accession: A69026
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-242 <MTH>
A;Cross-references: GB:AE000887; GB:AE000666; NID:g2622289; PIDN:AAB85781.1; PID:g2622290
A;Experimental source: strain Delta H
C;Genetics:
A;Gene: MTH1192
A;Start codon: GTG

Query Match 59.7%; Score 37; DB 2; Length 242;
Best Local Similarity 63.6%; Pred. No. 18;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 2 SFTLASSETGVG 13
||| | | :||
Db 113 SFTMVSPDTGV 123

RESULT 14
D95901
probable transcription regulator protein [imported] - Sinorhizobium meliloti
C;Species: Sinorhizobium meliloti
C;Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 14-Sep-2001
C;Accession: D95901
R;Finan, T.M.; Weidner, S.; Wong, K.; Buhrmester, J.; Chain, P.; Vorhölter, F.J.; Her
Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001
A;Title: The complete sequence of the 1,683 kb pSymB megaplasmid from the N2-fixing
A;Reference number: A95842; MUID:21396508; PMID:11481431
A;Accession: D95901
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-298 <KUR>
A;Cross-references: GB:AL591985; PIDN:CAC47087.1; PID:g13114...; GSPDB:GN00...
A;Experimental source: strain 1021, megaplasmid pSymB
R;Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubl
pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.
L.; Hyman, R.W.; Jones, T.
Science 293, 668-672, 2001
A;Authors: Kahn, D.; Kahn, M.L.; Kaiman, S.; Keating, D.H.; Kiss, E.; Krol, E.; Loria
hebault, P.; Vandenbol, M.; Vorhölter, F.J.; Weidner, S.; Weiss, S.; Wong, K.; Yeg
A;Title: The composite genome of the legume symbiont Sinorhizobium meliloti
A;Reference number: A96039; MUID:21368234; PMID:11474104
A;Contents: annotation
C;Genetics:
A;Gene: SMb20494
A;Genome: plasmid

Query Match 54.7%; Score 37; DB 2; Length 298;
Best Local Similarity 58.3%; Pred. No. 24;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 RSFTLASSETGV 12
|:|||| : |:
Db 17 RSFTLAAEDLGL 28

RESULT 15
F83895
metal tetracycline/H+ antiporter 'tetB [imported] - Bacillus halodurans (strain C-12
C;Species: Bacillus halodurans
C;Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001
C;Accession: F83895
R;Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fuji, F.; H
Nucleic Acids Res. 28, 4317-4331, 2000
A;Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans a
A;Reference number: A83650; MUID:20512582; PMID:11058142
A;Accession: F83895
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-459 <SI>
A;Cross-references: GB:AP001514; GB:BA000004; NID:g10174345; PIDN:BAB...; GSPDB:G...
A;Experimental source: strain C-125
C;Genetics:
A;Gene: tetB

Query Match 59.7%; Score 37; DB 2; Length 459;
Best Local Similarity 66.7%; Pred. No. 46;
Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 2 SFTLASSET:V3 1-
[illegible]
Db 621 SMTLAKKEE:V3 [illegible]

Search completed: August 26, 2002, 06:10:50
Job time: 111 sec

GenCore version 4.5
Copyright (c) 1993 - 2030 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 26, 2002, 06:13:25 ; Search time 139.28 Seconds
(without alignments)
10.367 Million cell updates/sec

Title: US-09-744-282-6
Perfect score: 62
Sequence: 1 RSPTLASSEPGVC 13

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747774 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747774

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_052802:*
1: /SIDS1/gcgdata/geneseq/geneseq-aa/AA_New:AA1985.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseq-aa/AA_New:AA1986.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseq-aa/AA_New:AA1987.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseq-aa/AA_New:AA1984.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseq-aa/AA_New:AA1985.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseq-aa/AA_New:AA1986.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseq-aa/AA_New:AA1987.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseq-aa/AA_New:AA1988.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseq-aa/AA_New:AA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseq-aa/AA_New:AA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseq-aa/AA_New:AA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseq-aa/AA_New:AA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseq-aa/AA_New:AA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseq-aa/AA_New:AA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseq-aa/AA_New:AA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseq-aa/AA_New:AA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseq-aa/AA_New:AA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseq-aa/AA_New:AA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseq-aa/AA_New:AA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseq-aa/AA_New:AA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseq-aa/AA_New:AA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseq-aa/AA_New:AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	62	100.0	13	17	AAR95076	Peptide conju. bf
2	62	100.0	16	19	AAW61597	Human cartilage gp
3	62	100.0	18	21	AAY69231	Peptide fragment o
4	62	100.0	18	21	AAY77401	Human cartilage gl
5	62	100.0	18	22	AAU03045	Human cartilage gp
6	62	100.0	18	22	AAU03046	Human cartilage gp
7	62	100.0	18	22	AAU03059	Human cartilage gp
8	62	100.0	18	22	AAU03065	Human cartilage gp
9	62	100.0	18	22	AAU03066	Human cartilage gp
10	62	100.0	18	22	AAU03067	Human cartilage gp
11	62	100.0	18	19	AAW61598	Human cartilage gp
12	62	100.0	16	18	AAW45140	Human cartilage gl
13	62	100.0	16	19	AAW61599	Human cartilage gp
14	62	100.0	105	22	AAU02074	Synthetic human ca
15	62	100.0	199	22	ABB56452	Human single chain
16	62	100.0	199	22	ABB56456	Human single chain
17	62	100.0	208	22	ABB56475	Human single chain
18	62	100.0	209	22	ABB56454	Human single chain
19	62	100.0	210	22	ABB56442	Human single chain
20	62	100.0	211	22	ABB56441	Human single chain
21	62	100.0	211	22	ABB56455	Human single chain
22	62	100.0	212	22	ABB56439	Human single chain
23	62	100.0	212	22	ABB56440	Human single chain
24	62	100.0	212	22	ABB56477	Human single chain
25	62	100.0	213	22	ABB56443	Human single chain
26	62	100.0	213	22	ABB56445	Human single chain
27	62	100.0	213	22	ABB56446	Human single chain
28	62	100.0	213	22	ABB56453	Human single chain
29	62	100.0	214	22	ABB56444	Human single chain
30	62	100.0	216	22	ABB56447	Human single chain
31	62	100.0	216	22	ABB56448	Human single chain
32	62	100.0	220	22	ABB56449	Human single chain
33	62	100.0	220	22	ABB56450	Human single chain
34	62	100.0	220	22	AAU02078	Synthetic human ca
35	62	100.0	281	22	AAU02079	Synthetic mutant h
36	62	100.0	450	22	AAU02076	Synthetic human mu
37	62	100.0	499	22	AAU02077	Synthetic multi hu
38	58	93.5	13	22	AAU03054	Human cartilage gp
39	58	93.5	13	22	AAU03073	Human cartilage gp
40	57	91.9	13	22	AAU03046	Human cartilage gp
41	57	91.9	13	22	AAU03047	Human cartilage gp
42	57	91.9	13	22	AAU03052	Human cartilage gp
43	57	91.9	13	22	AAU03057	Human cartilage gp
44	57	91.9	13	22	AAU03058	Human cartilage gp
45	57	91.9	13	22	AAU03062	Human cartilage gp

ALIGNMENTS

RESULT 1
AAR95076
ID AAR95076 standard; peptide; 13 AA.
XX
AC AAR95076;
XX
DT 17-DEC-1996 (first entry)
XX
DE Peptide conju. bf gp-39 autoantigen derived peptide (11).
XX
KW Human; cartilage glycoprotein-39; bf gp-39; autoantigen;
KW late-stage; rheumatoid arthritis; associated; destruction;
KW rheumatoid arthritis; disease model; diagnosis.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Peptide 1..11
FT /note= "bf gp-39 peptide (11)"
XX
PN WO9613517-A1.
XX
PD 09-MAY-1996.
XX
PF 25-OCT-1995; 95WO-EP04201.
XX
PR 07-APR-1995; 95EP-0200886.
PR 27-OCT-1994; 94EP-0203128.
XX
PA (AKZO) AKZO NOBEL NV.
XX
PI Boots AMH, Verheijden GFM;
XX
DR WPI; 1996-239447/24.

XX
PI Peptide(s) derived from human cartilage gp-39 autoantigen - used
PI to induce specific T cell tolerance, esp. for treatment of
PI rheumatoid arthritis, also useful in producing arthritic animal
PI models
XX
PS Claim 2; Page 28; 37pp; English.
XX
CC The present peptide, contg. a peptide derived from the human
CC cartilage glycoprotein 39 (hC gp-39) autoantigen, can be used to
CC induce specific T cell tolerance to hC gp-39 in a patient with
CC T cell mediated cartilage destruction, esp. rheumatoid arthritis
CC (RA). It may also be used to induce RA in mice, etc. (human disease
CC models), and to diagnose the presence of activated T-cells
CC autoreactive with hC gp-39, i.e. to identify patients suitable for
CC tolerising treatment.
CC Therapeutically the peptide is given in a dose of 0.1-100 ug/kg
CC by injection, or 10-50 microg/kg to produce a disease model. The
CC peptide has a specific tolerising effect on autoreactive T-cells,
CC but leaves other components of the immune system intact, and
CC therefore should have few side effects.
XX
SQ Sequence 13 AA;

Query Match 100.0%; Score 62; DB 17; Length 13;
Best Local Similarity 100.0%; Pred. No. 4.2e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KSFTLASSETGVG 13
||||||||||||| |
Db 1 rsftlassetqvq 13

RESULT 2
AAW61597
ID AAW61597 standard; protein; 13 AA.
XX
AC AAW61597;
XX
DT 27-OCT-1998 (first entry)
XX
DE Human Cartilage gp-39 epitope (i).
XX
KW Human; cartilage gp-39; monoclonal antibody; cell surface antigen;
KW rheumatoid arthritis; T-cell receptor.
XX
OS Homo sapiens.
XX
PN EP856520-A1.
XX
PD 05-AUG-1998.
XX
PF 02-DEC-1997; 97EP-0203769.
XX
PR 27-JUN-1997; 97EP-0201972.
PR 06-DEC-1996; 96EP-0203469.
XX
PA (ALKU) AKZO NOBEL NV.
PA (ALKU) AKZO NV.
XX
PI Steenbakkers PGA;
XX
DR WPI; 1998-401866/35.
XX
PT Preparation of monoclonal antibodies against cell surface antigens
PT used as diagnostic agents and to treat rheumatoid arthritis -
PT comprises immunising, then enriching, enriching and cloning B cells
PT and then selecting and cloning hybridoma
XX
PS Examples; Page 5; 24pp; English.
XX
CC The human cartilage gp-39 peptides AAW61597-W61599 were used in the
CC preparation of a monoclonal antibody against a cell surface ag. hC.
CC The antibody can be used for treatment of rheumatoid arthritis. It
CC is also useful against cell surface antigens and antibodies to T-cells
CC with the clonotypic structure of T cell receptors are used as diagnostic
CC reagents. The process allows antibodies to be raised against
CC antigens which are present in very low amounts or which are
CC conformationally unstable without the need to screen very large numbers
CC of hybridomas and without requiring high purity Antigen for immunisation.
XX
SQ Sequence 13 AA;

Query Match 100.0%; Score 62; DB 19; Length 13;
Best Local Similarity 100.0%; Pred. No. 4.2e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KSFTLASSETGVG 13
||||||||||||| |
Db 1 rsftlassetqvq 13

RESULT 3
AAY69231
ID AAY69231 standard; peptide; 13 AA.
XX
AC AAY69231;
XX
DT 30-MAY-2000 (first entry)
XX
DE Peptide fragment of human cartilage gp-39 (HC gp-39) protein.
XX
KW Human cartilage gp-39 protein; HC gp-39 protein; T cell activation;
KW inflammatory disease; lymphocyte; antigen; bystander suppression;
KW antigen-specific T-cell; tolerance; autoimmune disorder;
KW inflammatory disease; rheumatoid arthritis;
KW autoimmune inflammatory disease; autoantigen.
XX
OS Homo sapiens.
XX
PN WO200004917-A2.
XX
PD 03-FEB-2000.
XX
PF 19-JUL-1999; 99WO-EP05331.
XX
PR 23-JUL-1998; 98EP-0202471.
XX
PA (ALKU) AKZO NOBEL NV.
XX
PI Mittenburg AMM, Boots AMB;
XX
DR WPI; 2000-182537/16.
XX
PT Use of cartilage protein or peptides for treating inflammatory diseases,
PT especially rheumatoid arthritis
XX
PS Claim 2; Page 17; 29pp; English.
XX
CC AAY69226-38 represent fragments of the human cartilage gp-39 (HC gp-39)
CC protein. The peptides are T cell activators. The peptides are useful for
CC preventing inflammatory diseases, by modulating lymphocyte response
CC to antigens other than HC gp-39, present in the same tissue of
CC HC gp-39. By the induction of antigen specific T cell tolerance,
CC autoimmune disorders can be treated by bystander suppression. The
CC HC gp-39 peptides are useful for preventing inflammatory diseases,
CC especially rheumatoid arthritis. The fragments may also be used for
CC treating autoimmune inflammatory diseases. The HC gp-39 peptides are
CC especially beneficial in autoimmune conditions in which the
CC autoantigens associated with the condition are unknown.
XX
SQ Sequence 13 AA;

```
  Query Match             100.0%;  Score 62;  DB 21;  Length 13;
  Best Local Similarity   100.0%;  Prod. No. 4.2e-05;
  Matches   13;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 RSFTLASSETGVG 13
           |||||||||||||
Db       1 rsftlassetgvg 13


RESULT   4
AAY77401
ID   AAY77401 standard; peptide; 13 AA.
XX
AC   AAY77401;
XX
DT   22-MAY-2000  (first entry)
XX
DE   Human cartilage glycoprotein-39 (HC gp-39), residues 263-275.
XX
KW   Human cartilage glycoprotein-39; HC gp-39; YKL-40; epitope;
KW   autoreactive T-cell; tolerisation; desensitisation; autoantigen;
KW   rheumatoid arthritis; human chondrocyte protein; YKL-39;
KW   bystander suppression; autoimmune disorder; Graves' disease;
KW   primary glomerulonephritis; inflammatory bowel disease; diagnosis.
XX
OS   Homo sapiens.
XX
PN   WO200005254-A2.
XX
PD   03-FEB-2000.
XX
PF   16-JUL-1999;   99WO-EP005C.
XX
PR   23-JUL-1998;   98EP-0203470.
XX
PA   (AKZO ) AKZO NOBEL NV.
XX
PI   Verheijden GFM,  Boots AMH;
XX
DR   WPI; 2000-182443/16.
XX
PT   New T cell epitope peptides useful for inducing specific T cell
PT   tolerance to an autoantigen for treating autoimmune disorders,
PT   especially rheumatoid arthritis
XX
PS   Example 1; Page 10; 26pp; English.
XX
CC   The invention relates to peptides constituting amino acids 268-276 and
CC   amino acids 268-278 of human chondrocyte protein YKL-39
CC   (AAY77399-Y77401, respectively). The peptides of the invention may be
CC   used in pharmaceutical compositions for the induction of specific T-cell
CC   tolerance (desensitisation) to an autoantigen in patients
CC   suffering from autoimmune disorders, especially rheumatoid arthritis. A
CC   frequently found autoantigen in rheumatoid arthritis sufferers is human
CC   cartilage glycoprotein-39 (HC gp-39, also known as YKL-40). Several
CC   epitopes of HC gp-39 are recognised by peripheral blood T-cells from
CC   rheumatoid arthritis patients; one such epitope, HC gp-39 residues
CC   263-275, was recognised by T cells from 8 out of 18 patients tested.
CC   YKL-39 shares significant sequence identity to he gp-39. YKL-39 residues
CC   268-278 (AAY77401) constitutes a minimal epitope of HC gp-39 (263-275),
CC   and may be used for desensitisation of T cells with reactivity to HC
CC   gp-39 (263-275) in rheumatoid arthritis patients. Rheumatoid arthritis
CC   may be treated by the systemic administration of compositions comprising
CC   the peptides of the invention, optionally in combination with existing
CC   treatments. The peptides may also be used for modulating lymphocytes
CC   that are reactive to autoantigens other than YKL-39 and HC gp-39 but
CC   which are present in the same tissue as these autoantigens. By the
CC   induction of antigen specific T-cell tolerance, autoimmune disorders can
CC   be treated by bystander suppression. The conditions that may be treated
CC   via this approach include Graves' disease, primary glomerulonephritis
CC   and inflammatory bowel disease. The peptides, alone or with a detection
CC   agent, may additionally be used in a diagnostic composition. The peptides
CC   of the invention have a specific effect on autoreactive T-cells while
CC   leaving other components of the immune system intact. Therefore, the
CC   peptides should be safe, having no side effects. The present sequence
CC   represents residues 263-275 of human cartilage glycoprotein-39 (HC gp-39)
CC   which was used in an exemplification of the present invention.
XX
SQ   Sequence   13 AA;


  Query Match             100.0%;  Score 62;  DB 21;  Length 13;
  Best Local Similarity   100.0%;  Prod. No. 4.2e-05;
  Matches   13;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 RSFTLASSETGVG 13
           |||||||||||||
Db       1 rsftlassetgvg 13


RESULT   5
AAU03045
ID   AAU03045 standard; peptide; 13 AA.
XX
AC   AAU03045;
XX
DT   12-SEP-2001  (first entry)
XX
DE   Human cartilage gp-39 autoimmune peptide #1.
XX
KW   Human; autoimmune disease; epitope; antigen; rheumatoid arthritis;
KW   immunosuppression.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT                   /note "Deaminated"
XX
PN   WO200129081-A1.
XX
PD   26-APR-2001.
XX
PF   12-OCT-2000; 2000WO-EP10230.
XX
PR   18-OCT-1999;   99EP-0203427.
XX
PA   (AKZO ) AKZO NOBEL NV.
XX
PI   Van Staveren CJ,  Timmers CM,  Van Galen PJM,  Kneetel EMA,  Boots AMH;
PI   Miltenburg AMM;
XX
DR   WPI; 2001-367293/38.
XX
PT   Modified peptides derived from human cartilage gp-39 protein epitope
PT   are useful in immunotherapy, especially in the treatment of arthritis
PT   -
XX
PS   Claim 1; Page 36; 52pp; English.
XX
CC   The sequence represents Human cartilage gp-39 protein epitope (amino
CC   acids 263-275) which is used as the template for making modified
CC   peptides for use in immunotherapy. The peptides are useful in: induction
CC   tolerance induction in patients suffering from autoimmune diseases such
CC   as rheumatoid arthritis. The peptides provide antigen specific, non-toxic
CC   immunosuppression therapy as a very attractive alternative to
CC   non specific immunosuppression.
XX
SQ   Sequence   13 AA;


  Query Match             100.0%;  Score 62;  DB 22;  Length 13;
  Best Local Similarity   100.0%;  Prod. No. 4.2e-05;
  Matches   13;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 RSFTLASSETGVG 13
```

```
              ||||||||||||||
Db       1 rsttlassetqvg 13


RESULT   6
AAU03048
ID   AAU03048 standard; peptide; 13 AA.
XX
AC   AAU03048;
XX
DT   12-SEP-2001  (first entry)
XX
DE   Human cartilage gp-39 autoimmune peptide #4.
XX
KW   Human; autoimmune disease; epitope; antigen; rheumatoid arthritis;
KW   immunosuppression.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT                   /note= "Connected to CH3-(n-CH2-CH3)-n-CH2-(=O)"
FT   Modified-site   13
FT                   /note= "C-terminus is amidated"
XX
PN   WO200129081-A1.
XX
PD   26-APR-2001.
XX
PF   12-OCT-2000; 2000WO-EP10230.
XX
PR   18-OCT-1999;  99EP-0203427.
XX
PA   (AKZO ) AKZO NOBEL NV.
XX
PI   Van Staveren CJ,  Timmers CM,  Van Galen PJM,  Koentgen RMA,  Bloois AMH;
PI   Miltenburg AMM;
XX
DR   WPI; 2001-367293/38.
XX
PT   Modified peptides derived from human cartilage gp-39 protein epitope
PT   are useful in immunotherapy, especially in the treatment of arthritis
PT
XX
PS   Claim 12; Page 89; 52pp; English.
XX
CC   The sequence represents a modified peptide derived from Human cartilage
CC   gp-39 protein epitope (amino acids 263-275) for use in immunotherapy.
CC   The peptides are useful for inducing tolerance induction in patients
CC   suffering from autoimmune diseases such as rheumatoid arthritis.
CC   The peptides provide antigen specific, non-toxic immunosuppression
CC   therapy as a very attractive alternative to non-specific
CC   immunosuppression.
XX
SQ   Sequence   13 AA;


  Query Match             100.0%;  Score 62;  DB 22;  Length 13;
  Best Local Similarity   100.0%;  Pred. No. 4.2e-05;
  Matches   13;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 RSFTLASSETGVG 13
           |||||||||| ||
Db       1 rsttlassetqvg 13


RESULT   7
AAU03059
ID   AAU03059 standard; peptide; 13 AA.
XX
AC   AAU03059;
XX
```

```
DT   12-SEP-2001  (first entry)
XX
DE   Human cartilage gp-39 autoimmune peptide #15.
XX
KW   Human; autoimmune disease; epitope; antigen; rheumatoid arthritis;
KW   immunosuppression.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT                   /note= "N-terminal (N-methyl nicotinoyl)-"
XX
PN   WO200129081-A1.
XX
PD   26-APR-2001.
XX
PF   12-OCT-2000; 2000WO-EP10230.
XX
PR   18-OCT-1999;  99EP-0203427.
XX
PA   (AKZO ) AKZO NOBEL NV.
XX
PI   Van Staveren CJ,  Timmers CM,  Van Galen PJM,  Koentgen RMA,  Bloois AMH;
PI   Miltenburg AMM;
XX
DR   WPI; 2001-367293/38.
XX
PT   Modified peptides derived from human cartilage gp-39 protein epitope
PT   are useful in immunotherapy, especially in the treatment of arthritis
PT
XX
PS   Claim 12; Page 89; 52pp; English.
XX
CC   The sequence represents a modified peptide derived from Human cartilage
CC   gp-39 protein epitope (amino acids 263-275) for use in immunotherapy.
CC   The peptides are useful for inducing tolerance induction in patients
CC   suffering from autoimmune diseases such as rheumatoid arthritis.
CC   The peptides provide antigen specific, non-toxic immunosuppression
CC   therapy as a very attractive alternative to non-specific
CC   immunosuppression.
XX
SQ   Sequence   13 AA;


  Query Match             100.0%;  Score 62;  DB 22;  Length 13;
  Best Local Similarity   100.0%;  Pred. No. 4.2e-05;
  Matches   13;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 RSFTLASSETGVG 13
           |||||||||| ||||
Db       1 rsttlassetqvg 13


RESULT   8
AAU03065
ID   AAU03065 standard; peptide; 13 AA.
XX
AC   AAU03065;
XX
DT   12-SEP-2001  (first entry)
XX
DE   Human cartilage gp-39 autoimmune peptide #21.
XX
KW   Human; autoimmune disease; epitope; antigen; rheumatoid arthritis;
KW   immunosuppression.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
```

```
FT                   /note= "N-terminal acetyl"
FT   Misc-difference 1
FT                   /note= "Ile is a residue"
XX
PN   WO200129081-A1.
XX
PD   26-APR-2001.
XX
PF   12-OCT-2000; 2000WO-EP10230.
XX
PR   18-OCT-1999;   99EP-0203427.
XX
PA   (ALKU ) AKZO NOBEL NV.
XX
PI   Van Staveren CJ, Timmers CM, Van Galen PJM, Knegtel RMA, Boots AMH;
PI   Miltenburg AMM;
XX
DR   WPI; 2001-367293/38.
XX
PT   Modified peptides derived from human cartilage gp-39 protein epitope
PT   are useful in immunotherapy, especially in the treatment of arthritis
PT
XX
PS   Disclosure; Page 34; 52pp; English.
XX
CC   The sequence represents a modified peptide derived from Human cartilage
CC   gp-39 protein epitope (amino acids 263-275) for use in immunotherapy.
CC   The peptides are useful for inducing tolerance induction in patients
CC   suffering from autoimmune diseases such as rheumatoid arthritis.
CC   The peptides provide antigen-specific, non toxic immunosuppression
CC   therapy as a very attractive alternative to non-specific
CC   immunosuppression.
XX
SQ   Sequence   13 AA;


  Query Match             100.0%;  Score 62;  DB 22;  Length 13;
  Best Local Similarity   100.0%;  Pred. No. 4.2e-05;
  Matches   13;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 RSFTLASSEIVVG 13
           |||||||||||||
Db       1 rsftlassetvvg 13


RESULT   9
AAU03066
ID   AAU03066 standard; peptide; 13 AA.
XX
AC   AAU03066;
XX
DT   12-SEP-2001  (first entry)
XX
DE   Human cartilage gp-39 autoimmune peptide #22.
XX
KW   Human; autoimmune disease; epitope; antigen; rheumatoid arthritis;
KW   immunosuppression.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT                   /note= "N-terminal acetyl"
XX
PN   WO200129081-A1
XX
PD   26-APR-2001.
XX
PF   12-OCT-2000; 2000WO-EP10230.
XX
PR   18-OCT-1999;   99EP-0203427.
XX
```

```
PA   (ALKU ) AKZO NOBEL NV.
XX
PI   Van Staveren CJ, Timmers CM, Van Galen PJM, Knegtel RMA, Boots AMH;
PI   Miltenburg AMM;
XX
DR   WPI; 2001-367293/38.
XX
PT   Modified peptides derived from human cartilage gp-39 protein epitope
PT   are useful in immunotherapy, especially in the treatment of arthritis
PT
XX
PS   Example 15; Page 42; 52pp; English.
XX
CC   The sequence represents a modified peptide derived from Human cartilage
CC   gp-39 protein epitope (amino acids 263-275) for use in immunotherapy.
CC   The peptides are useful for inducing tolerance induction in patients
CC   suffering from autoimmune diseases such as rheumatoid arthritis.
CC   The peptides provide antigen specific, non toxic immunosuppression
CC   therapy as a very attractive alternative to non specific
CC   immunosuppression.
XX
SQ   Sequence   13 AA;


  Query Match             100.0%;  Score 62;  DB 22;  Length 13;
  Best Local Similarity   100.0%;  Pred. No. 4.2e-05;
  Matches   13;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 RSFTLASSEIVVG 13
           |||||||||||||
Db       1 rsftlassetvvg 13


RESULT   10
AAU03067
ID   AAU03067 standard; peptide; 13 AA.
XX
AC   AAU03067;
XX
DT   12-SEP-2001  (first entry)
XX
DE   Human cartilage gp-39 autoimmune peptide #23.
XX
KW   Human; autoimmune disease; epitope; antigen; rheumatoid arthritis;
KW   immunosuppression.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT                   /note= "N-terminal MeG-C(O)"
XX
PN   WO200129081-A1.
XX
PD   26-APR-2001.
XX
PF   12-OCT-2000; 2000WO-EP10230.
XX
PR   18-OCT-1999;   99EP-0203427.
XX
PA   (ALKU ) AKZO NOBEL NV.
XX
PI   Van Staveren CJ, Timmers CM, Van Galen PJM, Knegtel RMA, Boots AMH;
PI   Miltenburg AMM;
XX
DR   WPI; 2001-367293/38.
XX
PT   Modified peptides derived from human cartilage gp-39 protein epitope
PT   are useful in immunotherapy, especially in the treatment of arthritis
PT
XX
PS   Example 9; Page 25; 52pp; English.
```

```
XX
CC   The sequence represents a modified peptide derived from human cartilage
CC   gp 39 protein epitope (amino acids 264-275) for use in immunotherapy.
CC   The peptides are useful for inducing tolerance induction in patients
CC   suffering from autoimmune diseases such as rheumatoid arthritis.
CC   The peptides provide antigen specific, non-toxic immunosuppressive
CC   therapy as a very attractive alternative to non specific
CC   immunosuppression..
XX
SQ   Sequence   15 AA;


  Query Match             100.0%;  Score 62;  DB 22;  Length 15;
  Best Local Similarity   100.0%;  Pred. No. 4.2e-05;
  Matches   13;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 RSFTLASSEIGVG 13
           |||||||||||||
Db       3 rsftlassetgvg 15


RESULT  11
AAW61598
ID   AAW61598 standard; protein; 15 AA.
XX
AC   AAW61598;
XX
DT   27-OCT-1998  (first entry)
XX
DE   Human Cartilage gp 39 epitope (ii).
XX
KW   Human; cartilage gp 39; monoclonal antibody; cell surface antigen;
KW   rheumatoid arthritis; T cell receptor.
XX
OS   Homo sapiens.
XX
PN   EP856520-A1.
XX
PD   05-AUG-1998.
XX
PF   02-DEC-1997;  97EP-0203760.
XX
PR   27-JUN-1997;  97EP-0201972.
PR   06-DEC-1996;  96EP-0203465.
XX
PA   (ALKO ) AKZO N BEL NV.
PA   (ALKO ) AKZO NV.
XX
PI   Steenbakkers PGA;
XX
DR   WPI; 1998-400806/35.
XX
PT   Preparation of monoclonal antibodies against cell surface antigens
PT   used as diagnostic agents and to treat rheumatoid arthritis -
PT   comprises immunising, then enriching, enriching and cloning B cells
PT   and then selecting and cloning hybridoma
XX
PS   Examples; Page 13; 23pp; English.
XX
CC   The human cartilage gp 39 peptides AAW61597-W61600 were used in the
CC   preparation of a monoclonal antibody against a cell surface antigen.
CC   The antibody can be used for treatment of rheumatoid arthritis. It
CC   is also useful against cell surface antigens and antibodies reactive
CC   with the clonotypic structure of T cell receptors are used as diagnostic
CC   reagents. The process allows antibodies to be raised against
CC   antigens which are present in very low amounts or which are
CC   conformationally unstable without the need to screen very large numbers
CC   of hybridomas and without requiring high purity Antigen for enrichment.
XX
SQ   Sequence   15 AA;


  Query Match             100.0%;  Score 62;  DB 19;  Length 15;
  Best Local Similarity   100.0%;  Pred. No. 4.2e-05;
  Matches   13;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 RSFTLASSEIGVG 13
           |||||||||||||
Db       3 rsftlassetgvg 15


RESULT  12
AAW45140
ID   AAW45140 standard; peptide; 15 AA.
XX
AC   AAW45140;
XX
DT   28-APR-1998  (first entry)
XX
DE   Human cartilage glycoprotein 39 derived peptide #41.
XX
KW   Articular cartilage; immunosuppressive therapy; antibody; T cell; B cell;
KW   immunological tolerance; T cell; human cartilage glycoprotein 39;
KW   HC gp 39; rheumatoid arthritis; epitope.
XX
OS   Synthetic.
OS   Homo sapiens.
XX
PN   WO9740068-A1.
XX
PD   30-OCT-1997.
XX
PF   22-APR-1997;  97WO-EP02051.
XX
PR   24-APR-1996;  96EP-0201119.
XX
PA   (ALKO ) AKZO NOBEL NV.
XX
PI   Boots AMH;  Verheijden JPM;
XX
DR   WPI; 1997-535775/49.
XX
PT   Peptide suitable for use in antigen specific immunosuppressive
PT   therapy - resembles or mimics epitope present on HC gp 39 -
PT   inducing systemic immunological tolerance to rheumatoid arthritis
PT   autoantigen
XX
PS   Disclosure; Page 16; 82pp; English.
XX
CC   The present sequence represents a peptide which resembles or mimics an
CC   epitope present in human cartilage glycoprotein 39 (HC gp 39), an
CC   autoantigen in rheumatoid arthritis. The invention is directed to a
CC   consisting of 10-25 amino acid residues comprising at least 10 to 15
CC   following 19 sequences: LVCYYISWS; FLTIHEIYS; TIYSIANES; IRCIIVYI;
CC   FIKSVIIEL; FLGLLAWLL; LRSSRRRGG; YGIAKIGLH; LHFICIMIY; FICIMIYIL;
CC   FRGQEDASP; YAVCAMLGL; MLSCAPAS; LAYYLICFI; LRGAIVHIT; YIFFIGIA;
CC   LAGAMVWAL; VWALDLDDE; or EICIAFGGS. They can be used preferably in
CC   antigen specific immunosuppressive therapy, particularly the treatment
CC   of T-cell mediated destruction of articular cartilage in autoimmune
CC   diseases (e.g. rheumatoid arthritis). They can also be used to induce
CC   activated autoreactive T cells in an individual. The peptides have a
CC   specific effect on the autoreactive T cells, thus leaving the other
CC   components of the immune system intact, unlike the non specific
CC   suppressive effect of immunosuppressive drugs, and do not cause severe
CC   side effects. The peptides are predominantly recognised by activated T
CC   T cells from rheumatoid arthritis patients, but rarely by T cells from
CC   healthy donors.
XX
SQ   Sequence   16 AA;


  Query Match             100.0%;  Score 62;  DB 19;  Length 16;
  Best Local Similarity   100.0%;  Pred. No. 4.2e-05;
  Matches   13;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 RSFTLASSEIGVG 13
```

```
                                  1
Db       2 rsftlassetqvq 14


RESULT  13
AAW61599
ID   AAW61599 standard; protein; 15 AA.
XX
AC   AAW61599;
XX
DT   27-OCT-1998  (first entry)
XX
DE   Human Cartilage gp-39 epitope (III).
XX
KW   Human; cartilage gp-39; monoclonal antibody; cell surface antigen;
KW   rheumatoid arthritis; T-cell receptor.
XX
OS   Homo sapiens.
XX
PN   EP856520-A1.
XX
PD   05-AUG-1998.
XX
PF   02-DEC-1997;   97EP-0203769.
XX
PR   27-JUN-1997;   97EP-0201972.
PR   06-DEC-1996;   96EP-0203405.
XX
PA   (SOLV ) SOLVAY & CIE SA NV.
XX
PI   Stoenbakkers PCA;
XX
DR   WPI; 1998-400805/34.
XX
PT   Preparation of monoclonal antibodies against cell surface antigens
PT   used as diagnostic agents and to treat rheumatoid arthritis
PT   - comprises immunising, then enriching, enriching and cloning B cells
PT   and then selecting and cloning hybridoma
XX
PS   Examples; Page 34; 23pp; English.
XX
CC   The human cartilage gp-39 peptides AAW61597-AAW61599 were used in the
CC   preparation of a monoclonal antibody against a cell surface antigen.
CC   The antibody can be used for treatment of rheumatoid arthritis. It
CC   is also useful against cell surface antigens and antibodies reactive
CC   with the clono-typic structure of T-cell receptors are used as diagnostic
CC   reagents. The process allows antibodies to be raised against
CC   antigens which are present in very low amounts or which are
CC   conformational/unstable without the need to screen very large numbers
CC   of hybridomas and without requiring high purity Antigen for enrichment.
XX
SQ   Sequence   15 AA;


  Query Match             100.0%;  Score 62;  DB 19;  Length 15;
  Best Local Similarity   100.0%;  Pred. No. 5.2e-05;
  Matches   13;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 RSFTLASSETGVG 13
           |||||||||||||
Db       2 rsftlassetqvq 14


RESULT  14
AAB62074
ID   AAB62074 standard; Protein; 105 AA.
XX
AC   AAB62074;
XX
DT   07-SEP-2001  (first entry)
XX
DE   Synthetic human target autoantigen shCLP/RA.
XX
KW   shCLP/RA; immunogenic epitope cluster; IEC;
KW   synthetic human chondrocyte glycoprotein 69/cartilage link protein;
KW   autoantigen; autoimmune disease; multiple sclerosis;
KW   rheumatoid arthritis, insulin dependent diabetes mellitus;
KW   myasthenia gravis; uveitis; autoimmune hepatitis; thyroiditis; insulitis;
KW   orchitis; idiopathic thrombocytopenic purpura; inflammatory disease;
KW   Crohn's disease; ulcerative colitis.
XX
OS   Synthetic.
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 103
FT                   /label= OTHER
FT                   /note= "Other= STOP"
XX
PN   WO200131037-A2.
XX
PD   03-MAY-2001.
XX
PF   26-OCT-2000; 2000WO-IL00688.
XX
PR   27-OCT-1999;   99IL-0132611.
XX
PA   (YEDA ) YEDA RES & DEV CO LTD.
XX
PI   Ben-Nun A,  Kerlero De Rosbo N,  Sappler GP;
XX
DR   WPI; 2001-308515/31.
DR   N-PSDB; AAS04827.
XX
PT   Novel synthetic human target autoantigen gene useful for treating
PT   autoimmune diseases such as multiple sclerosis, insulin-dependent
PT   diabetes mellitus, rheumatoid arthritis, myasthenia gravis, and uveitis
PT   -
XX
PS   Claim 173; Fig 51; 182pp; English.
XX
CC   The sequence represents shCLP/RA (synthetic human chondrocyte
CC   glycoprotein 69/cartilage link protein) containing immunogenic epitope
CC   clusters (IEC) from chondrocyte glycoprotein 69 and cartilage link
CC   protein. The synthetic human target autoantigen genes of the invention
CC   comprise sequences coding for at least 2 IEC's of autoantigen(s) related
CC   to a specific autoimmune disease. The synthetic human target autoantigen
CC   genes are useful for treating autoimmune diseases such as multiple
CC   sclerosis, insulin-dependent diabetes mellitus, rheumatoid arthritis,
CC   myasthenia gravis, uveitis, autoimmune hepatitis, thyroiditis, insulitis,
CC   orchitis, idiopathic thrombocytopenic purpura, and inflammatory diseases
CC   (Crohn's disease, ulcerative colitis). The synthetic human target
CC   autoantigen genes are also useful for diagnosis and/or monitoring the
CC   progression of the autoimmune disease.
XX
SQ   Sequence   105 AA;


  Query Match             100.0%;  Score 62;  DB 22;  Length 105;
  Best Local Similarity   100.0%;  Pred. No. 0.00044;
  Matches   13;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 RSFTLASSETGVG 13
           |||||||||||||
Db      29 rsftlassetqvq 41


RESULT  15
ABB56452
ID   ABB56452 standard; Protein; 199 AA.
XX
AC   ABB56452;
XX
DT   25-FEB-2002  (first entry)
XX
```

```
DE   Human single chain MHC class II molecule C0528 AC
XX
KW   Human; MHC; major histocompatibility complex; MHC class II; multimer;
KW   single chain; immunosuppressive; antidiabetic; anti-inflammatory;
KW   antianaemic; antirheumatoid; antiarthritic; neuroprotective; vaccine;
KW   autoimmune disease; insulin dependent diabetes; multiple sclerosis;
KW   myasthenia gravis; pernicious anaemia; autoimmune encephalomyelitis;
KW   rheumatoid arthritis; systemic lupus erythematosus.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
PN   WO200170,45 A1.
XX
PD   27 SEP 2001.
XX
PF   22 MAR 2001; 2001WO-US09616.
XX
PR   22 MAR 2000; 2000US-191274P.
PR   15 MAY 2000; 2000US-204249P.
PR   23 JAN 2001; 2001US-264004P.
XX
PA   (CORI ) CORIXA CORP.
XX
PI   Carter D;  Zhu S;  Arimilli S;  Wang A;
XX
DR   WPI; 2001-616471/71.
DR   N-PSDB; AB199022.
XX
PT   Multimeric complex for treating autoimmune diseases, comprises first
PT   and second single chain MHC class II molecules, each comprising alphal
PT   and betal domain linked through amino acid linker and multimerization
PT   domain
XX
PS   Claim 41; Page 89; 147pp; English.
XX
CC   The invention relates to a multimeric complex comprising a first
CC   recombinant single chain major histocompatibility complex (MHC) class
CC   II molecule and a second recombinant single chain MHC class II molecule,
CC   each comprising an alphal domain and a betal domain linked through an
CC   amino acid linker and a multimerisation domain. The first and the
CC   second molecule are linked through the multimerisation domain to form
CC   a multimeric complex. The complex is useful for treating autoimmune
CC   diseases. It is useful for treating insulin dependent diabetes,
CC   multiple sclerosis, myasthenia gravis, pernicious anaemia, autoimmune
CC   encephalomyelitis (EAE), rheumatoid arthritis and systemic lupus
CC   erythematosus. The present sequence is a single chain MHC class II
CC   molecule of the invention.
XX
SQ   Sequence   199 AA;


  Query Match             100.0%;  Score 62;  DB 22;  Length 199;
  Best Local Similarity   100.0%;  Pred. No. 0.000098;
  Matches   14;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 RSFFLASSEIGVG 14
             ||||| ||| |||
Db       6 RSFFLASSETGVG 18



Search completed: August 26, 2002, 06:13:25
Job time: 267 sec
```

Gen' ce version 4.5
Copyright (c) 1993 - 2002 Compugen Ltd.

GM protein - protein seq vs. protein sw model

Run on: August 26, 2002, 06:12:05 ; Search time 55.56 Seconds
(without alignments)
13,986 Million cell updates/sec

Title: US-09-744-282-6
Perfect score: 62
Sequence: 1 PSEPLASSEPGVG 13

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38714550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0
Maximum DB seq length: 2300000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: SwissProt_40:*

Prob. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	62	100	383	1	CH3L1_HUMAN	P36222 homo sapien
2	46	74.2	390	1	CH3L2_HUMAN	q15782 homo sapien
3	43	69.4	381	2	CH3L1_MOUSE	q61362 mus musculu
4	39	62.9	387	1	REGN_MYCTU	o33197 mycobacteri
5	38	61.3	387	1	RBGK_MOUSE	P97329 mus musculu
6	37	59.7	390	1	RBGK_HUMAN	o95239 homo sapien
7	35	56.5	187	1	Y054_MIBPI	P47706 bacteriopha
8	35	56.5	383	1	MBH_MYCLE	P46362 mycobacteri
9	35	56.5	402	1	CNTB_SCHPO	q09722 schizosacch
10	35	56.5	445	1	FLGL_VIBPA	q03435 vibrio para
11	35	56.5	573	1	REGN_TREPA	O83456 treponema p
12	35	56.5	686	2	HMT1_HELPJ	q92152 helicobacte
13	35	56.5	689	2	HMT1_HELPY	q50415 helicobacte
14	34	54.8	194	1	ENTB_BACHD	q48547 bacillus ha
15	34	54.8	210	1	FCL1_CHICK	P31999 gallus gall
16	34	54.8	410	2	CBP3_AQUAE	o66676 aquifex aeo
17	34	54.8	647	1	YIE1_YEAST	P38389 saccharomyc
18	34	54.8	782	2	SMIP_MOUSE	q62179 mus musculu
19	34	54.8	991	2	FHAB_BORPE	P31487 bordetella
20	33	53.2	96	1	CRAA_ECOLI	P37107 escherichia
21	33	53.2	294	1	YTFI_BACAN	q81dz2 bacillus an
22	33	53.2	360	1	YAPK_BACSU	P37473 bacillus su
23	33	53.2	366	1	YCGI_ECOLI	P76015 escherichia
24	33	53.2	449	1	IRA8_HUMAN	q96a65 homo sapien
25	33	53.2	449	1	IRA8_MOUSE	q9jhi2 mus musculu
26	33	53.2	467	1	EHEE_ECOLI	P77445 escherichia
27	33	53.2	467	1	EHEE_SALTY	P41793 salmonella
28	33	53.2	572	1	REGN_STRCO	q9x220 streptomyce
29	33	53.2	576	1	REGN_BACSU	P17864 bacillus su
30	33	53.2	587	1	REGN_MYCLE	q49878 mycobacteri
31	33	53.2	625	1	HTPG_BACHD	q9k651 bacillus ha
32	33	53.2	626	1	HTPG_BACSU	P54208 bacillus su
33	33	53.2	1076	1	NOP1_YEAST	P20676 saccharomyc
34	32	51.6	104	1	YBK0_YEAST	P38815 saccharomyc
35	32	51.6	125	1	RL19_TREPA	O83887 treponema p
36	32	51.6	147	1	YUXK_BACSU	P40761 bacillus su
37	32	51.6	158	1	RISB_NEIMA	Q9jqv6 neisseria m
38	32	51.6	218	1	YU27_STRCO	Q9rkm5 streptomyce
39	32	51.6	224	1	TCA2_VIBCH	P29024 vibrio cho
40	32	51.6	228	1	EFIB_CRYSA	q43758 cryptospir
41	32	51.6	242	1	BIOD_BACHD	Q9kex0 bacill.us h
42	32	51.6	266	1	INBA_MYCAV	o07400 mycobacteri
43	32	51.6	272	1	NIKC_ECOLI	P33932 escherichia
44	32	51.6	300	1	NAHK_PSEPU	P11434 pseudomonas
45	32	51.6	321	1	ISP_BACCS	P29144 bacillus st

ALIGNMENTS

```
RESULT 1
CH3L1_HUMAN
ID   CH3L1_HUMAN     STANDARD;      PRT:   383 AA.
AC   P36222; P30923;
DT   01-JUL-1993 (Rel. 26, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Chitinase-3 like protein 1 precursor (Cartilage glycoprotein-39) (GP
DE   39) (39 kDa synovial protein) (YKL-40).
GN   CHI3L1.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC   TISSUE=Cartilage;
RX   MEDLINE=94064658; PubMed=8245017;
RA   Hakala B.E., White C., Recklies A.D.;
RT   "Human cartilage gp-39, a major secretory product of articular
RT   chondrocytes and synovial cells, is a mammalian member of a chitinase
RT   protein family.";
RL   J. Biol. Chem. 268:25803-25810(1993).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Blood;
RX   MEDLINE=97386591; PubMed=9244440;
RA   Rehli M., Krause S.W., Andreesen R.;
RT   "Molecular characterization of the gene for human cartilage gp-39
RT   (CHI3L1), a member of the chitinase protein family and marker for
RT   late stages of macrophage differentiation.";
RL   Genomics 43:221-225(1997).
RN   [3]
RP   SEQUENCE OF 22-45.
RX   MEDLINE=93328983; PubMed=2475759;
RA   Nyirkos P., Golds E.E.;
RT   "Human synovial cells secrete a 39 kDa protein similar to a bovine
RT   mammary protein expressed during the non-lactating period.";
RL   Biochem. J. 269:265-268(1990).
CC   -!- FUNCTION: MAY PLAY AN IMPORTANT ROLE IN THE CAPACITY OF CELLS TO
CC       RESPOND TO AND COPE WITH CHANGES IN THEIR ENVIRONMENT.
CC   -!- SUBCELLULAR LOCATION: Extracellular.
CC   -!- TISSUE SPECIFICITY: PRESENT IN ARTICULAR CHONDROCYTES, SYN VIAL
CC       CELLS AS WELL AS IN LIVER. EXPRESSED IN MUSCLE TISSUES, LUNG,
CC       PANCREAS, NON-NUCLEAR CELLS, OR FIBROBLASTS.
CC   -!- PTM: GLYCOSYLATED.
CC   -!- SIMILARITY: BELONGS TO FAMILY 18 OF GLYCOSYL HYDROLASES.
CC   -----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----------------------------------------------------------------------
```

```
DR   EMBL; M80629; AAA16074.1;  -.
DR   EMBL; Y08374; CAA69661.1;  -.
DR   EMBL; Y08375; CAA69661.1; JOINED.
DR   EMBL; Y08376; CAA69661.1; JOINED.
DR   EMBL; Y08377; CAA69661.1; JOINED.
DR   EMBL; Y08378; CAA69661.1; JOINED.
DR   PIR; S10677; S10677.
DR   PIR; A33162; A33162.
DR   MIM; 601526; -.
DR   InterPro; IPR001579; Chitinase_2.
DR   InterPro; IPR001223; Glyco_hydro_18.
DR   Pfam; PF00704; Glyco_hydro_18; 1.
DR   PROSITE; PS01095; CHITINASE_18; FALSE_NEG.
KW   Glycoprotein; Signal.
FT   SIGNAL          1     27
FT   CHAIN          28    383       CHITINASE 3 LIKE PROTEIN 2.
FT   CARBOHYD       69     69       N-LINKED (GLCNAC...) (POTENTIAL).
SQ   SEQUENCE   383 AA;  42613 MW;  76ADD8258EEEC2D3 CRC64;


  Query Match             100.0%;  Score 62;  DB 1;  Length 383;
  Best Local Similarity   100.0%;  Pred. No. 0.00032;
  Matches   13;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 RSFTLASSEFGVG 13
           ||||||||||||| ||
Db     269 RSFTLASSEFGVG 281


RESULT   2
CH3L2_HUMAN
ID   CH3L2_HUMAN     STANDARD;      PRT:   390 AA.
AC   Q15782; Q15783; Q15749;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Chitinase-3-like protein 2 precursor (YKL-39) (Chondrocyte protein
DE   39).
GN   CHI3L2.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A. (SHORT AND LONG ISOFORMS).
RA   Grossman A., Matsuyama T., Baker E., Waterhouse P., Sutherland G.R.,
RA   Mak T.W.;
RT   "Cloning of a novel lymphoid restricted human chit nase and
RT   localization to 1p13.3.";
RL   Submitted (MAY-1996) to EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A. (SHORT ISOFORM).
RC   TISSUE=Articular cartilage;
RX   MEDLINE=99325075; PubMed=10206729;
RA   Hu B., Trinh K., Figueira W.F., Price P.A.;
RT   "Isolation and sequence of a novel human chondrocyte protein related
RT   to mammalian members of the chitinase protein family.";
RL   J. Biol. Chem. 271:19415-19420(1996).
CC   -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS; A SHORT FORM (SHORT HERE) AND A
CC       LONG FORM MAY BE PRODUCED BY ALTERNATIVE SPLICING;
CC   -!- TISSUE SPECIFICITY: HIGHEST EXPRESSION IN CHONDROCYTES, FOLLOWED
CC       BY SYNOVIOCYTES, LUNG AND HEART. NOT DETECTED IN BRAIN, SPLEEN,
CC       PANCREAS, AND LIVER.
CC   -!- SIMILARITY: BELONGS TO FAMILY 18 OF GLYCOSYL HYDROLASES.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; D58515; AAB04534.1;  -.
DR   EMBL; D58514; AAB04533.1;  -.
DR   EMBL; U49845; AAC50597.1; ALT_INIT.
DR   MIM; 601526; -.
DR   InterPro; IPR001579; Chit_base_2.
DR   InterPro; IPR001223; Glyco_hydro_18.
DR   Pfam; PF00704; Glyco_hydro_18; 1.
DR   PROSITE; PS01095; CHITINASE_18; FALSE_NEG.
KW   Hydrolase; Signal; Alternative splicing.
FT   SIGNAL          1     26       POTENTIAL.
FT   CHAIN          27    390       CHITINASE 3-LIKE PROTEIN 2.
FT   CARBOHYD       35     35       N-LINKED (GLCNAC...) (POTENTIA
FT   VARSPLIC        1     12       MGATIMGQKSLW -> SQQTWDGGQGEVWEEVVVQHHEE
FT                                  AIIWDSKIEPGSIIDSEHPDSI (IN ISOFORM 2).
SQ   SEQUENCE   390 AA;  43506 MW;  97B05A2F4AA3D7A7 CRC64;


  Query Match             74.2%;  Score 46;  DB 1;  Length 390;
  Best Local Similarity   83.3%;  Pred. No. 0.25;
  Matches   10;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       2 SFTLASSEFGVG 13
           |||||:|:|:||
Db     272 SFTLASAFTIVG 284


RESULT   3
CH3L_MOUSE
ID   CH3L_MOUSE     STANDARD;      PRT:   381 AA.
AC   Q61362;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Chitinase-3 like protein 1 precursor (Cartilage glycoprotein-39)
DE   (GP-39) (BRP39 protein).
GN   CHI3L1 OR BRP39.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BALB/c; TISSUE=Breast;
RX   MEDLINE=95060797; PubMed=7527540;
RA   Morrison B.W., Leder P.;
RT   "neu and ras initiate murine mammary tumors that share genetic markers
RT   generally absent in c-myc and int-2 initiated tumors.";
RL   Oncogene 9:3417-3426(1994).
CC   -!- FUNCTION: MAY PLAY AN IMPORTANT ROLE IN TISSUE REMODELING DURING
CC       BRAIN GROWTH AND/OR IN THE RESPONSE TO INJURY (BY
CC       SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: Extracellular (By similarity).
CC   -!- SIMILARITY: BELONGS TO FAMILY 18 OF GLYCOSYL HYDROLASES.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; X92305; CAA63603.1;  -.
DR   MGD; MGI:1340899; Chi3l1.
DR   InterPro; IPR001579; Chit_base_2.
DR   InterPro; IPR001223; Glyco_hydro_18.
DR   Pfam; PF00704; Glyco_hydro_18; 1.
DR   ProSite; PS01095; CHITINASE_18; FALSE_NEG.
KW   Glycoprotein; Signal.
FT   SIGNAL          1     21       POTENTIAL.
FT   CHAIN          22    381       CHITINASE 3-LIKE PROTEIN 1.
FT   CARBOHYD       60     60       N-LINKED (GLCNAC...) (POTENTIA
SQ   SEQUENCE   381 AA;  43021 MW;  EF05B1B2D94F945C CRC64;
```

```
  Query Match             69.4%;  Score 43;  DB 1;  Length 981;
  Best Local Similarity   69.2%;  Pred. No. 0.93;
  Matches    9;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 RSFILASSEIGVG 13
           :|||:|||| :|
Db     264 KSFILASSENQLG 276


RESULT   4
RECN_MYCTU
ID   RECN_MYCTU     STANDARD;      PRT:   587 AA.
AC   O53197;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel.  40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   DNA repair protein recN (Recombination protein N).
GN   RECN OR RV1896c OR MTCY25.18.
OS   Mycobacterium tuberculosis.
OC   Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC   Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX   NCBI_TaxID=1773;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=H37RV;
RX   MEDLINE=98206096; PubMed=9634230;
RA   Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA   Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekaia F.,
RA   Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA   Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
RA   Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
RA   Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA   Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA   Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
RT   "Deciphering the biology of Mycobacterium tuberculosis from the
RT   complete genome sequence.";
RL   Nature 393:537-544(1998).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CIP 105511 / Oshkosh;
RA   Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA   Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
RA   Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
RA   Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
RA   Bishai W.,
RT   "Whole genome comparison of Mycobacterium tuberculosis clinical and
RT   laboratory strains.";
RL   Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: MAY BE INVOLVED IN RECOMBINATIONAL REPAIR OF DAMAGED
CC       DNA (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE RECN FAMILY.
CC   -----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----------------------------------------------------------------------
DR   EMBL; Z98268; CAB10453.1; -.
DR   EMBL; AE007035; AAK46004.1; -.
DR   HSSP; M11745;
DR   TubercuList; Rv1896c; -.
DR   InterPro; IPR003439; ABC_transporter;.
KW   DNA repair; ATP-binding; Complete proteome.
FT   NP_BIND      25     30       ATP (POTENTIAL).
FT   CONFLICT    410    410       E -> G (IN REF. 2).
SQ   SEQUENCE   587 AA;  62228 MW;  4B705EEB60102E7 CRC64;
```

```
  Query Match             62.9%;  Score 39;  DB 1;  Length 58;
  Best Local Similarity   58.8%;  Pred. No. 8.8;
  Matches    7;  Conservative    2;  Mismatches    4;  Indels    1;  Gaps    1;

Qy       1 RSFILASSETGVG 13
           | ||:  : ||| |
Db      22 RGFTVLIGETGIG 34


RESULT   5
RB6K_MOUSE
ID   RB6K_MOUSE     STANDARD;      PRT:   887 AA.
AC   P97429;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Rab6 kinesin (RAB6 interacting kinesin like protein) (Kinesin-like
DE   protein 174).
GN   RAB6KIFL.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BALB/C; TISSUE=testis;
RX   MEDLINE=98101856; PubMed=9438865;
RA   Echard A., Jollivet F., Martinez O., Lacapere J.-J., Rousselet A.,
RA   Janoueix-Lerosey I., Goud B.;
RT   "Interaction of a Golgi-associated kinesin-like protein with Rab6.";
RL   Science 279:580-585(1998).
CC   -!- FUNCTION: INTERACTS WITH GUANOSINE TRIPHOSPHATE (GTP) BOUND FORMS
CC       OF RAB6A AND RAB6B. MAY ACT AS A MOTOR REQUIRED FOR THE RETROGRADE
CC       RAB6 REGULATED TRANSPORT OF GOLGI MEMBRANES AND ASSOCIATED
CC       VESICLES ALONG MICROTUBULES. HAS A MICROTUBULE PLUS END DIRECTED
CC       MOTILITY.
CC   -!- SUBCELLULAR LOCATION: GOLGI.
CC   -!- TISSUE SPECIFICITY: UBIQUITOUSLY EXPRESSED, WITH HIGHEST LEVELS IN
CC       SPLEEN AND TESTIS.
CC   -!- SIMILARITY: BELONGS TO THE KINESIN LIKE PROTEIN FAMILY.
CC   -----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----------------------------------------------------------------------
DR   EMBL; Y09642; CAA70845.1; -.
DR   HSSP; P17119; 3KAR.
DR   MGD; MGI:1201682; Rab6kifl.
DR   InterPro; IPR001752; kinesin.
DR   Pfam; PF00225; kinesin; 1.
DR   PRINTS; PR00380; KINESINHEAVY.
DR   SMART; SM00129; KISc; 1.
DR   PROSITE; PS00411; KINESIN_MOTOR_1@MAIN1; FALSE_NEG
DR   PROSITE; PS50067; KINESIN_MOTOR_2@MAIN2; 1.
KW   Motor protein; Microtubules; ATP-binding; Coiled coil; Golgi stack;
KW   Protein transport.
FT   DOMAIN       69    494       KINESIN-MOTOR.
FT   NP_BIND     159    166       ATP (POTENTIAL).
FT   DOMAIN      559    804       COILED COIL (POTENTIAL).
FT   DOMAIN      805    887       GLOBULAR (POTENTIAL).
SQ   SEQUENCE   887 AA;  99879 MW;  7143CCB261A4EA15 CRC64;
```

```
  Query Match             61.3%;  Score 38;  DB 1;  Length 88;
  Best Local Similarity   66.7%;  Pred. No. 22;
  Matches    8;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       2 SFILASSEIGVG 13
           || | :|: |||
```

```
Db     104 SFALKSNEKGVG 114


RESULT   6
RB6K_HUMAN
ID   RB6K_HUMAN      STANDARD;      PRT:   890 AA.
AC   O95235;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Rab6 kinesin 6 (RAB6 interacting kinesin like protein) (KIF20_2).
GN   RAB6KIFL.
OS   Homo sapiens (human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Endothelial cells;
RX   MEDLINE=99252600; PubMed=10233894;
RA   Borrevoets A.J.J., Foutlin R.D., van Zonneveld A.-J., de Vries C.J.M.,
RA   ten Cate J.W., Pannekoek H.;
RT   "Vascular endothelial genes that are responsive to tumor necrosis
RT   factor-alpha in vitro are expressed in atherosclerotic lesions,
RT   including inhibitor of apoptosis protein-1, stannin, and two novel
RT   genes.";
RL   Blood 93:3418-3431(1999).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Brain;
RX   MEDLINE=20267854; PubMed=10806357;
RA   Lai F., Fernald A.A., Zhao N., Le Beau M.M.;
RT   "cDNA cloning, expression pattern, genomic structure and chromosomal
RT   location of RAB6KIFL, a human kinesin-like gene.";
RL   Gene 248:117-125(2000).
CC   -!- FUNCTION: INTERACTS WITH GUANOSINE TRIPHOSPHATE (GTP)-BOUND FORMS
CC       OF RAB6A AND RAB6B. MAY ACT AS A MOTOR REQUIRED FOR THE RETROGRADE
CC       RAB6-REGULATED TRANSPORT OF GOLGI MEMBRANES AND ASSOCIATED
CC       VESICLES ALONG MICROTUBULES. HAS A MICROTUBULE PLUS END-DIRECTED
CC       MOTILITY (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: GOLGI (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE KINESIN-LIKE PROTEIN FAMILY.
CC   ---------- ---- ----- ----- ----- ----- ----- ----- ----- ----- ----- --
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------- ---- ----- ----- ----- ----- ----- ----- ----- ----- ----- --
DR   EMBL; AF070672; AAC83230.1; -.
DR   EMBL; AF153329; AAD37806.1; -.
DR   HSSP; P17119; 3KAR.
DR   MIM; 605664; .
DR   InterPro; IPR001752; Kinesin.
DR   Pfam; PF00225; kinesin; 2.
DR   PRINTS; PR00380; KINESINHEAVY.
DR   SMART; SM00129; KISc; 1.
DR   PROSITE; PS00411; KINESIN_MOTOR_DOMAIN; FALSE_NEG.
DR   PROSITE; PS50067; KINESIN_MOTOR_DOMAIN2; 1.
KW   Motor protein; Microtubules; ATP-binding; Coiled coil; Golgi stack;
KW   Protein transport.
FT   DOMAIN       73     405       KINESIN-MOTOR.
FT   NP_BIND      104    107       ATP (POTENTIAL).
FT   DOMAIN       611    762       COILED COIL (POTENTIAL).
FT   DOMAIN       765    890       GLOBULAR (POTENTIAL).
SQ   SEQUENCE   890 AA;  100277 MW;  66202461549605] CRC64;


  Query Match             59.7%;  Score 47;  DB 1;  Length 890;
  Best Local Similarity   58.8%;  Pred. No. 44;
  Matches    7;  Conservative    2;  Mismatches    4;  Indels    0;  Gaps    1;

QY       2 SFILASSEIGVG 13
           | | | ||| |
Db     134 SFALKSNEKGIG 145


RESULT   7
YO34_BPBP1
ID   YO34_BPBP1      STANDARD;      PRT:   187 AA.
AC   P51738;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Hypothetical 20.8 kDa protein in LYS 3'region (ORF4).
OS   Bacteriophage BP1.
OC   Viruses; dsDNA viruses, no RNA stage; Caudovirales; Myoviridae.
OX   NCBI_TaxID=10690;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BP1C1;
RX   MEDLINE=96279738; PubMed=8702706;
RA   Esposito D., Fitzmaurice W.P., Benjamin R.C., Goodman S.D.;
RA   Waldman A.S., Scocca J.J.;
RT   "The complete nucleotide sequence of bacteriophage BP1 DNA.";
RL   Nucleic Acids Res. 24:2360-2368(1996).
CC   ---------- ---- ----- ----- ----- ----- ----- ----- ----- ----- ----- --
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------- ---- ----- ----- ----- ----- ----- ----- ----- ----- ----- --
DR   EMBL; U24159; AAB09221.1; -.
KW   Hypothetical protein.
SQ   SEQUENCE   187 AA;  20761 MW;  A0A679D7161A702E CRC64;


  Query Match             56.9%;  Score 45;  DB 1;  Length 187;
  Best Local Similarity   63.6%;  Pred. No. 15;
  Matches    7;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

QY       3 FILASSEIGVG 13
           | || :|||
Db      72 FNLAEAEIGKG 82


RESULT   8
MRP_MYCLE
ID   MRP_MYCLE       STANDARD;      PRT:   383 AA.
AC   P53382;
DT   01-OCT-1996 (Rel. 34, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Mrp protein homolog.
GN   MRP OR ML1080.
OS   Mycobacterium leprae.
OC   Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC   Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX   NCBI_TaxID=1769;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Smith D.R., Robison K.;
RL   Submitted (MAY-1996) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=TN;
RX   MEDLINE=21128732; PubMed=11234002;
RA   Cole S.T., Eiglmeier K., Parkhill J., James K.D., Thomson N.R.,
RA   Wheeler P.R., Honore N., Garnier T., Churcher C., Harris D.,
RA   Mungall K., Basham D., Brown D., Chillingworth T., Connor R.,
RA   Davies R.M., Devlin K., Duthoy S., Feltwell T., Fraser A., Hance N.,
```

```
RA   Bilbrey S., Bristol J., Jones K., Lacey K., Mitchum J., Moore J.,
RA   Murphy L., Oliver S., Quail M.A., Rajandream M.A., Rutherford K.M.,
RA   Rutter S., Seeger K., Simon S., Simmonds M., Skelton J., Squares R.,
RA   Squares S., Stevens K., Taylor K., Whitehead S., Woodward J.R.,
RA   Barrell B.G.;
RT   "Massive gene decay in the leprosy bacillus.";
RL   Nature 409:1007-1011(2001).
CC   !  FUNCTION: NOT KNOWN.
CC   !  SIMILARITY: BELONGS TO THE MSH2/MSH6 FAMILY. PLAY FUNCTION
CC       PROTEINS.
CC   -!- -- ------- ---------- -- ------- --- - -- - -- -- -- -
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/
CC   or send an email to license@isb-sib.ch).
CC   -- - ------- ---------- -- ------- --- - -- - -- -- -- -
DR   EMBL; U15180; AAA27009.1; ALT_INIT.
DR   EMBL; AL583922; CAC43451.1; -.
DR   Leproma; ML0840; -.
DR   InterPro; IPR002743; HDF59.
DR   InterPro; IPR007695; Mrp.
DR   Pfam; PF11083; DUF59; 1.
DR   Prodom; PD005595; DUF59; 1.
DR   PROSITE; PS01215; MRP; 1.
KW   ATP binding; Complete proteome.
FT   NP BIND      1.      144      ATP (POTENTIAL).
SQ   SEQUENCE   383 AA;  43643 MW;  435B8F29A8C8279E CRC64;


  Query Match             56.5%;  Score 45;  DB 1;  Length 383;
  Best Local Similarity   58.8%;  Pred. No. 43;
  Matches    7;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 PSPTLASSEIGVG 13
           | : :|| | :||
Db     120 RVYIVASGKAGVG 132


RESULT   9
ONPB_SCHPO
ID   ONPB_SCHPO     STANDARD;      PRT:   432 AA.
AC   Q09722; Q40S29
DT   01-NOV-1995 (Rel. 32, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Cdc9/Fen1 endonuclease/exonuclease/5'-3' ...
GN   ONPB OR SPAC57A7.01 OR SPAC1F5.04;
OS   Schizosaccharomyces pombe (Fission yeast).
OC   Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC   Schizosaccharomycetales; Schizosaccharomycetaceae;
OC   Schizosaccharomyces
OX   NCBI_TaxID=4896;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=972;
RA   Zimmermann W., Wambutt R., Mebourrill R.C., Rajandream M.A.,
RA   Barrell B.G.;
RL   Submitted (NOV-1995) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE OF 3-40, FROM N.A.
RC   STRAIN=972;
RA   Doelin K., Thuriaux P., Barrell B.G., Rajandream M.A., Walsh S.V.;
RL   Submitted (SEP-1996) to the EMBL/GenBank/DDBJ databases.
CC   !  SIMILARITY: CONTAINS 1 PCI DOMAIN.
CC   -!- -- ------- ---------- -- ------- --- - -- - -- -- -- -
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/
CC   or send an email to license@isb-sib.ch).
CC   -- - ------- ---------- -- ------- --- - -- - -- -- -- -
DR   EMBL; AL133154; CAB61449.1; -.
DR   EMBL; Z50113; CAA90459.1; -.
DR   InterPro; IPR000717; PCI.
DR   SMART; SM00088; PINI; 1.
KW   Signalosome.
FT   DOMAIN      274     365       PCI.
SQ   SEQUENCE   402 AA;  45086 MW;  CD972E78BBFC0A3C CRC64;


  Query Match             56.5%;  Score 45;  DB 1;  Length 402;
  Best Local Similarity   70.0%;  Pred. No. 45;
  Matches    7;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       2 SPTLASSEIGVG 11
           |:|:|||:| |
Db     482 SPTIASSKKGG 492


RESULT   10
FLBL_VIBPA
ID   FLBL_VIBPA     STANDARD;      PRT:   445 AA.
AC   Q03475;
DT   01-OCT-1993 (Rel. 27, Created)
DT   01-OCT-1993 (Rel. 27, Last sequence update)
DT   15-OCT-2001 (Rel. 40, Last annotation update)
DE   Lateral flagellar hook associated protein 2 (HAP2, (Filament cap
DE   protein) (Flagellar cap protein).
GN   FLBL OR LAFB.
OS   Vibrio parahaemolyticus.
OC   Bacteria; Proteobacteria; Gamma subdivision; Vibrionaceae; Vibrio.
OX   NCBI_TaxID=670;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BB22;
RX   MEDLINE; 93273702; PubMed=8501040;
RA   McCarter L.L., Wright M.E.;
RT   "Identification of genes encoding components of the swarmer cell
RT   flagellar motor and propeller and a sigma factor controlling
RT   differentiation of Vibrio parahaemolyticus.";
RL   J. Bacteriol. 175:3361-3371(1993).
CC   !  FUNCTION: REQUIRED FOR THE MORPHOGENESIS AND FOR THE ELONGATION OF
CC      THE FLAGELLAR FILAMENT BY FACILITATING POLYMERIZATION OF THE
CC      FLAGELLIN MONOMERS AT THE TIP OF GROWING FILAMENT. FORMS A CAPPING
CC      STRUCTURE, WHICH PREVENTS FLAGELLIN SUBUNITS (TRANSPORTED THROUGH
CC      THE CENTRAL CHANNEL OF THE FLAGELLUM) FROM LEAKING OUTWARD
CC      POLYMERIZATION AT THE DISTAL END. ESSENTIAL FOR SWARMING MOTILITY.
CC   !  SUBUNIT: HOMOPENTAMER (BY SIMILARITY).
CC   !  SUBCELLULAR LOCATION: FLAGELLAR.
CC   !  MISCELLANEOUS: V.PARAHAEMOLYTICUS POSSESSES TWO FLAGELLAR SYSTEMS:
CC      A SINGLE POLAR FLAGELLUM PROPELS THE BACTERIUM IN LIQUID
CC      (SWIMMING), WHILE MULTIPLE LATERAL (PERITRICHOUS) FLAGELLA MOVE
CC      THE BACTERIUM OVER SURFACES (SWARMING). THE POLAR FLAGELLUM IS
CC      SYNTHESIZED CONSTITUTIVELY BUT LATERAL FLAGELLA ARE EXPRESSED ONLY
CC      UNDER CONDITIONS IN WHICH THE POLAR FLAGELLUM IS NOT FUNCTIONAL.
CC   !  SIMILARITY: BELONGS TO THE FLID FAMILY.
CC   -!- -- ------- ---------- -- ------- --- - -- - -- -- -- -
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/
CC   or send an email to license@isb-sib.ch).
CC   -- - ------- ---------- -- ------- --- - -- - -- -- -- -
DR   EMBL; L06176; AAA27530.1; -.
DR   EMBL; U52957; AAB07351.1; -.
DR   PIR; B40590; B40590.
DR   InterPro; IPR003481; FliD.
DR   Pfam; PF02465; FliD; 1.
KW   Flagella; Coiled coil.
```

```
FT   DOMAIN       388    424       COILED COIL (POTENTIAL).
SQ   SEQUENCE   445 AA;  48264 MW;  0F7E7089A525852A CRC64;


  Query Match             56.5%;  Score 35;  DB 1;  Length 445;
  Best Local Similarity   70.0%;  Pred. No. 39;
  Matches    7;  Conservative    2;  Mismatches    1;  Indels    1;  Gaps    0;

Qy       3 FTLASSEIGV 12
           | | |:|||
Db     177 FMLSSIEIGV 186


RESULT  11
REC'N_TREPA
ID   REC'N_TREPA     STANDARD;      PRT:   573 AA.
AC   O83456;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   DNA repair protein recN (Recombination protein N).
GN   REC'N OR TP0442.
OS   Treponema pallidum.
OC   Bacteria; Spirochaetales; Spirochaetaceae; Treponema.
OX   NCBI_TaxID=16;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=NICHOLS;
RX   MEDLINE=98332770; PubMed=9665876;
RA   Fraser C.M., Norris S.J., Weinstock G.M., White O., Sutton G.G.,
RA   Dodson R., Gwinn M., Hickey E.K., Clayton R., Ketchum K.A.,
RA   Sodergren E., Hardham J.M., McLeod M.P., Salzberg S., Peterson J.,
RA   Khalak H., Richardson D., Howell J.K., Chidambaram M., Utterback T.,
RA   McDonald L., Artiach P., Bowman C., Cotton M.D., Fujii C., Garland S.,
RA   Hatch B., Horst K., Roberts K., Sandusky M., Weidman J., Smith H.O.,
RA   Venter J.C.;
RT   "Complete genome sequence of Treponema pallidum, the syphilis
RT   spirochete.";
RL   Science 281:375-388(1998).
CC   -!- FUNCTION: MAY BE INVOLVED IN RECOMBINATIONAL REPAIR OF DAMAGED
CC       DNA (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE REC'N FAMILY.
CC   ---------- ---------- ---------- ---------- ---------- ----------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------- ---------- ---------- ---------- ---------- ----------
DR   EMBL; AE001221; AAC65429.1; -.
DR   TIGR; TP0442; -.
DR   InterPro; IPR003439; ABC_transporter.
KW   DNA repair; ATP binding; Complete proteome.
FT   NP_BIND      29     36       ATP (POTENTIAL).
SQ   SEQUENCE   573 AA;  63571 MW;  94388258901926E3 CRC64;


  Query Match             56.5%;  Score 35;  DB 1;  Length 573;
  Best Local Similarity   63.6%;  Pred. No. 51;
  Matches    7;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       3 FTLASSEIGVG 13
           | : . :| |
Db      24 FIALSGEIGAG 34


RESULT  12
HMCT_HELPJ*
ID   HMCT_HELPJ     STANDARD;      PRT:   686 AA.
AC   Q92L53;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
DE   Cadmium, zinc and cobalt transporting ATPase (EC 3.6.3.-).
DE   (EC 3.6.3.5).
GN   CADA OR JHP0727
OS   Helicobacter pylori J99 (Campylobacter pylori J99).
OC   Bacteria; Proteobacteria; epsilon subdivision; Helicobacter et rel.
OC   Helicobacter.
OX   NCBI_TaxID=85963;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=99120557; PubMed=9882662;
RA   Alm R.A., Ling L.-S.L., Moir D.T., King B.L., Brown E.D., Doig P.C.,
RA   Smith D.R., Noonan B., Guild B.C., deJonge B.L., Carmel G.,
RA   Tummino P.J., Caruso A., Uria-Nickelsen M., Mills D.M., Ives C.,
RA   Gibson R., Merberg D., Mills S.D., Jiang Q., Taylor D.E., Vovis G.F.,
RA   Trust T.J.;
RT   "Genomic-sequence comparison of two unrelated isolates of the human
RT   gastric pathogen Helicobacter pylori.";
RL   Nature 397:176-180(1999).
CC   -!- FUNCTION: COUPLES THE HYDROLYSIS OF ATP WITH THE TRANSPORT OF
CC       CADMIUM, ZINC AND COBALT OUT OF THE CELL.
CC   -!- CATALYTIC ACTIVITY: ATP + H(2)O + Cd(2+)(IN) = ADP + PHOSPHATE +
CC       CD(2+)(OUT).
CC   -!- CATALYTIC ACTIVITY: ATP + H(2)O + ZN(2+)(IN) = ADP + PHOSPHATE +
CC       ZN(2+)(OUT).
CC   -!- SUBCELLULAR LOCATION: Internal membrane protein.
CC   -!- SIMILARITY: BELONGS TO THE CATION TRANSPORT ATPASE FAMILY
CC       (E1-E2 ATPASES) SUBFAMILY IB.
CC   -!- SIMILARITY: CONTAINS 2 HMA DOMAIN.
CC   ---------- ---------- ---------- ---------- ---------- ----------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------- ---------- ---------- ---------- ---------- ----------
DR   EMBL; AE001514; AAD06808.1; -.
DR   InterPro; IPR001696; Cad_ATPase.
DR   InterPro; IPR001757; E1_E2_ATPase.
DR   InterPro; IPR001934; HMA.
DR   InterPro; IPR001454; Hydrolase.
DR   Pfam; PF00122; E1-E2_ATPase; 1.
DR   Pfam; PF00403; HMA; 1.
DR   Pfam; PF00702; Hydrolase; 1.
DR   PRINTS; PR00119; CATATPASE.
DR   PROSITE; PS00154; ATPASE_E1_E2; 1.
DR   PROSITE; PS01047; HMA_1; 1.
DR   PROSITE; PS50846; HMA_2; 1.
KW   Hydrolase; Transmembrane; Phosphorylation; Membrane; ATP-binding;
KW   Metal-binding; Cadmium; Zinc; Cobalt; Complete proteome.
FT   DOMAIN       1     72       CYTOPLASMIC (BY SIMILARITY).
FT   TRANSMEM     73    92       1 (BY SIMILARITY).
FT   DOMAIN       93    102      EXTRACELLULAR (BY SIMILARITY).
FT   TRANSMEM     103   124      2 (BY SIMILARITY).
FT   DOMAIN       125   131      CYTOPLASMIC (BY SIMILARITY).
FT   TRANSMEM     132   151      3 (BY SIMILARITY).
FT   DOMAIN       152   154      EXTRACELLULAR (BY SIMILARITY).
FT   TRANSMEM     155   174      4 (BY SIMILARITY).
FT   DOMAIN       175   308      CYTOPLASMIC (BY SIMILARITY).
FT   TRANSMEM     309   327      5 (BY SIMILARITY).
FT   DOMAIN       328   332      EXTRACELLULAR (BY SIMILARITY).
FT   TRANSMEM     333   351      6 (BY SIMILARITY).
FT   DOMAIN       351   635      CYTOPLASMIC (BY SIMILARITY).
FT   TRANSMEM     636   657      7 (BY SIMILARITY).
FT   DOMAIN       658   665      EXTRACELLULAR (BY SIMILARITY).
FT   TRANSMEM     666   681      8 (BY SIMILARITY).
FT   DOMAIN       682   686      CYTOPLASMIC (BY SIMILARITY).
FT   DOMAIN         1    63       HMA.
FT   MOD_RES      388   388      PHOSPHORYLATION (BY SIMILARITY).
```

```
FT   METAL*        11     11       POTENTIAL.
FT   METAL         14     14       POTENTIAL.
FT   METAL        583    583       MAGNESIUM (BY SIMILARITY).
FT   METAL        587    587       MAGNESIUM (BY SIMILARITY).
SQ   SEQUENCE   686 AA;  74837 MW;  01E69A3B057476E2 CRC64;


  Query Match             54.5%;  Score 1.5;  DB 1;  Length 686;
  Best Local Similarity   60.0%;  Pred. No. 0.2;
  Matches    6;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       4 ILASSEISVG  8
           |||:|:|:|
Db     590 ILASADVGIG 599


RESULT  14
HMC1_HELPY
ID   HMC1_HELPY     STANDARD;      PRT;   686 AA.
AC   O25465;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
DE   Cadmium, zinc and cobalt transporting ATPase (EC 3.6.3.3)
DE   (EC 3.6.3.6).
GN   CADA OR HP0791.
OS   Helicobacter pylori (Campylobacter pylori).
OC   Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC   Helicobacter.
OX   NCBI_TaxID=210;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=69A;
RX   MEDLINE=96142943; PubMed=8550603;
RA   Melchers K., Weitzenegger T., Buhmann A., Steinhilber W., Sachs G.,
RA   Schaefer K.P.;
RT   "Cloning and membrane topology of a P type ATPase from Helicobacter
RT   pylori.";
RL   J. Biol. Chem. 271:416-457(1996).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=26695 / ATCC 700392;
RX   MEDLINE=97394467; PubMed=9252185;
RA   Tomb J.-F., White O., Kerlavage A.R., Clayton R.A., Sutton G.G.,
RA   Fleischmann R.D., Ketchum K.A., Klenk H.-P., Gill S., Dougherty B.A.,
RA   Nelson K., Quackenbush J., Zhou L., Kirkness E.F., Peterson S.,
RA   Loftus B., Richardson D., Dodson R., Khalak H.G., Glodek A.,
RA   McKenney K., Fitzegerald L.M., Lee N., Adams M.D., Hickey E.K.,
RA   Berg D.E., Gocayne J.D., Utterback T.R., Peterson J.D., Kelley J.M.,
RA   Cotton M.D., Weidman J.M., Fujii C., Bowman C., Watthey L., Wallin E.,
RA   Hayes W.S., Borodovsky M., Karp P.D., Smith H.O., Fraser C.M.,
RA   Venter J.C.;
RT   "The complete genome sequence of the gastric pathogen Helicobacter
RT   pylori.";
RL   Nature 388:539-547(1997).
RN   [3]
RP   TOPOLOGY.
RC   STRAIN=69A;
RX   MEDLINE=99348838; PubMed=10417744;
RA   Herrmann L., Seowan D., Gartner R., Mobley H.L., Haas R.,
RA   Schaefer K.P., Melchers K.;
RT   "Helicobacter pylori cadA encodes an essential Cd(II)-Zn(II)-Co(II)
RT   resistance factor influencing urease activity.";
RL   Mol. Microbiol. 33:524-536(1999).
RN   [4]
RP   SUBSTRATE SPECIFICITY.
RC   STRAIN=69A;
RX   MEDLINE=20041191; PubMed=11078781;
RA   Melchers K., Schmidgall A., Buhmann A., Weitzenegger T., Belli R.,
RA   Gross S., Ehrmann M.;
RT   "Membrane topology of CadA homologous P-type ATPase of Helicobacter
RT   pylori as determined by expression of phoA fusions in Escherichia
RT   coli and the positive inside rule.";
RL   Res. Microbiol. 150:507-520(1999).
CC   -!- FUNCTION: COUPLES THE HYDROLYSIS OF ATP WITH THE TRANSPORT OF
CC       CADMIUM, ZINC AND COBALT OUT OF THE CELL.
CC   -!- CATALYTIC ACTIVITY: ATP + H(2)O + CD(2+)(IN)   ADP + PHOSPHATE +
CC       CD(2+)(OUT).
CC   -!- CATALYTIC ACTIVITY: ATP + H(2)O + ZN(2+)(IN)   ADP + PHOSPHATE +
CC       ZN(2+)(OUT).
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein
CC   -!- SIMILARITY: BELONGS TO THE CATION TRANSPORT ATPASES FAMILY
CC       (E1-E2 ATPASES), SUBFAMILY IB.
CC   -!- SIMILARITY: CONTAINS 1 HMA DOMAIN.
CC   -!- -------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -------------------------------------------------------------------------
DR   EMBL; L46864; AAA93044.1; -.
DR   EMBL; AE000591; AAD07839.1; -.
DR   PIR; H70791; H70791.
DR   InterPro; IPR001757; E1-E2_ATPase.
DR   InterPro; IPR001994; HMA.
DR   InterPro; IPR001454; Hydrolase.
DR   Pfam; PF00122; E1-E2_ATPase; 1.
DR   Pfam; PF00403; HMA; 1.
DR   Pfam; PF00702; Hydrolase; 1.
DR   PROSITE; PS00154; ATPASE_E1_E2; 1.
DR   PROSITE; PS01047; HMA_1; 1.
DR   PROSITE; PS50846; HMA_2; 1.
KW   Hydrolase; Transmembrane; Phosphorylation; Magnesium; ATP-binding;
KW   Metal-binding; Cadmium; Zinc; Cobalt; Complete proteome.
FT   DOMAIN        1     72       CYTOPLASMIC.
FT   TRANSMEM     73     92       1.
FT   DOMAIN       93    102       EXTRACELLULAR.
FT   TRANSMEM    103    124       2.
FT   DOMAIN       125    131       CYTOPLASMIC.
FT   TRANSMEM     132    151       3.
FT   DOMAIN       152    154       EXTRACELLULAR.
FT   TRANSMEM     155    174       4.
FT   DOMAIN       175    308       CYTOPLASMIC.
FT   TRANSMEM     309    327       5.
FT   DOMAIN       328    332       EXTRACELLULAR.
FT   TRANSMEM     333    350       6.
FT   DOMAIN       351    635       CYTOPLASMIC.
FT   TRANSMEM     636    657       7.
FT   DOMAIN       658    665       EXTRACELLULAR.
FT   TRANSMEM     666    681       8.
FT   DOMAIN       682    686       CYTOPLASMIC.
FT   DOMAIN         1     63       HMA.
FT   MOD_RES      488    488       ASPARTYL PHOSPHATE (BY SIMILARITY).
FT   METAL         11     11       POTENTIAL.
FT   METAL         14     14       POTENTIAL.
FT   METAL        583    583       MAGNESIUM (BY SIMILARITY).
FT   METAL        587    587       MAGNESIUM (BY SIMILARITY).
FT   VARIANT       24     24       E -> K (IN STRAIN 69A).
FT   VARIANT       39     39       K -> R (IN STRAIN 69A).
FT   VARIANT       69     69       I -> A (IN STRAIN 69A).
FT   VARIANT       80     80       I -> V (IN STRAIN 69A).
FT   VARIANT       83     83       M -> A (IN STRAIN 69A).
FT   VARIANT      148    148       F -> C (IN STRAIN 69A).
FT   VARIANT      173    174       VS -> IA (IN STRAIN 69A).
FT   VARIANT      198    198       E -> A (IN STRAIN 69A).
FT   VARIANT      209    209       V -> I (IN STRAIN 69A).
FT   VARIANT      228    228       V -> I (IN STRAIN 69A).
FT   VARIANT      251    251       N -> S (IN STRAIN 69A).
FT   VARIANT      345    345       K -> E (IN STRAIN 69A).
FT   VARIANT      456    456       V -> L (IN STRAIN 69A).
FT   VARIANT      504    504       V -> I (IN STRAIN 69A).
FT   VARIANT      526    526       V -> A (IN STRAIN 69A).
FT   VARIANT      555    555       H -> Y (IN STRAIN 69A).
```

```
SQ   SEQUENCE   686 AA;  75012 MW;  8777A49DB0697E1C CRC64;


  Query Match             56.5%;  Score 35;  DB 1;  Length 686;
  Best Local Similarity   60.0%;  Pred. No. 62;
  Matches    6;  Conservative    3;  Mismatches    1;  Indels    2;  Gaps    0;

Qy       4 TLASSEIGVG 13
           ||:|:::.:.
Db     500 TLASAIVGIG 509


RESULT  14
ENGB_BACHD
ID   ENGB_BACHD     STANDARD;      PRT;   194 AA.
AC   Q9KRF7;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Probable GTP-binding protein EngB.
GN   ENGB OR BH3049.
OS   Bacillus halodurans.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group;
OC   Bacillus/Staphylococcus group; Bacillus.
OX   NCBI_TaxID=86665;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C-125 / JCM 9153;
RX   MEDLINE=20512582; PubMed=11058132;
RA   Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R.,
RA   Fuji F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,
RA   Horikoshi K.;
RT   "Complete genome sequence of the alkaliphilic bacterium Bacillus
RT   halodurans and genomic sequence comparison with Bacillus subtilis.";
RL   Nucleic Acids Res. 28:4317-4331(2000).
CC   -!- FUNCTION: NECESSARY FOR NORMAL CELL DIVISION AND FOR SOME
CC       MAINTENANCE OF NORMAL SEPTATION (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE ENGB FAMILY.
CC   -----------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----------------------------------------------------------------
DR   EMBL; AP001517; BAB06768.1; -.
DR   InterPro; IPR001806; Ras_transformed.
DR   PRINTS; PR00449; RASTRNSFRMNG.
KW   Cell division; Septation; GTP-binding; Complete proteome.
FT   NP_BIND      30     37       GTP (POTENTIAL).
FT   NP_BIND      75     79       GTP (POTENTIAL).
FT   NP_BIND     142    145       GTP (POTENTIAL).
SQ   SEQUENCE   194 AA;  21924 MW;  F810A7994AD70FB8 CRC64;


  Query Match             54.8%;  Score 34;  DB 1;  Length 194;
  Best Local Similarity   100.0%;  Pred. No. 24;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       7 SSEIGVG 13
           |||||||
Db     175 SSEIGVG 181


RESULT  15
POLI_CHICK
ID   POLI_CHICK     STANDARD;      PRT;   246 AA.
AC   P13450;
DT   01-MAR-1989 (Rel. 10, Created)
DT   01-MAR-1989 (Rel. 10, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Retrovirus-related POL polyprotein [Contains: Reverse transcriptase
DE   (EC 2.7.7.49); Endonuclease] (Fragments).
GN   POL.
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=9031;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=85270471; PubMed=2410912;
RA   Dunwiddie C., Faras A.J.;
RT   "Presence of retrovirus reverse transcriptase related gene sequences
RT   in avian cells lacking endogenous avian leukosis viruses.";
RL   Proc. Natl. Acad. Sci. U.S.A. 82:5097-5101(1985).
RN   [2]
RP   REVISION TO 2.
RA   Dunwiddie C., Faras A.J.;
RL   Submitted (OCT-1988) to the EMBL/Genbank/DDBJ databases.
CC   -!- ---------- .
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----------------------------------------------------------------
DR   EMBL; K03271; -; NOT_ANNOTATED_CDS.
DR   EMBL; K02572; -; NOT_ANNOTATED_CDS.
DR   PIR; A25279; A25279.
DR   Pfam; PF00665; rve; 1.
KW   Hydrolase; Transferase; RNA-directed DNA polymerase; Nucleotidyltransferase;
KW   Endonuclease; Polyprotein.
FT   NON_TER       1      1
FT   NON_CONS    111    112
FT   NON_TER     246    246
SQ   SEQUENCE   246 AA;  26854 MW;  EA036E7C07927C76 CRC64;


  Query Match             54.8%;  Score 34;  DB 1;  Length 246;
  Best Local Similarity   72.7%;  Pred. No. 42;
  Matches    8;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       3 FTIASSEIGVG 13
           || ||| | :|
Db     198 FILASSSIGQG 208


Search completed: August 26, 2002, 06:29:55
Job time: 931 sec
```



GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

SW protein protein search, using sw model

Run on: August 26, 2002, 06:23:08 ; Search time 13.75 Seconds
(without alignments)
17.274 Million cell updates/sec

Title: US-09-744-282-6
Perfect score: 62
Sequence: 1 RSFTLASSETGVG 13

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 5621.2 seqs, 1729042.0 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_19:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_plant:*
10: sp_rodent:*
11: sp_virus:*
12: sp_vertebrate:*
13: sp_unclassified:*
14: sp_virus:*
15: sp_bacteriap:*
16: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	62	100	383	4	Q96HI7	Q96hi7 homo sapien
2	51	82.3	383	6	Q29411	Q29411 sus scrofa
3	48	77.4	382	6	Q18949	Q18949 bos taurus
4	46	74.2	352	11	Q9WTV1	Q9wtv1 rattus norv
5	46	74.2	390	4	Q96EP7	Q96ep7 homo sapien
6	43.5	70.2	396	11	Q9D7Q1	Q9d7q1 mus musculu
7	43	69.4	381	11	Q99J84	Q99j84 mus musculu
8	41	66.1	807	13	Q9XED2	Q9xed2 trifolium ag
9	41	66.1	829	13	Q9XED3	Q9xed3 acetilops ta
10	41	66.1	830	1	Q40074	Q40074 triticum ae
11	40.5	65.3	387	4	Q9H3V8	Q9h3v8 homo sapien
12	40.5	65.3	466	4	Q18231	Q18231 homo sapien
13	40	64.5	185	2	Q38187	Q38187 bacteriophi
14	38	61.3	261	13	Q9HBA0	Q9hba0 halobacteri
15	38	61.3	441	2	O52483	O52483 serratia ma
16	38	61.3	531	1	Q9SRK9	Q9srk9 arabidopsis
17	38	61.3	867	5	Q9V8J3	Q9v8j3 dl septa
18	37	59.7	242	17	O27260	O27260 methanother
19	37	59.7	298	16	Q92W69	Q92w69 rhizobium m
20	37	59.7	459	16	Q9KB62	Q9kb62 bacillus ha
21	37	59.7	793	12	Q99CY1	Q99cy1 bovine bet
22	36	58.1	251	17	O27403	O27403 methanother
23	36	58.1	266	16	O69626	O69626 mycobacteri
24	36	58.1	278	2	Q50282	Q50282 mycobacteri
25	36	58.1	298	2	Q92406	Q92406 pseudomonas
26	36	58.1	451	16	Q9CCX7	Q9ccx7 mycobacteri
27	36	58.1	424	16	Q9PPC8	Q9ppc8 campylobact
28	36	58.1	446	10	Q9SXN2	Q9sxn2 pyrus pyrif
29	36	58.1	648	5	Q9VRU7	Q9vru7 drosophila
30	36	58.1	650	5	Q9UIJ7	Q9uij7 drosophila
31	36	58.1	664	2	Q9RHF5	Q9rhf5 salmonella
32	36	58.1	709	2	Q93BG7	Q93bg7 clostridium
33	36	58.1	810	10	Q9XGB3	Q9xgb3 triticum ae
34	36	58.1	833	10	Q9XGB1	Q9xgb1 triticum ae
35	36	58.1	833	10	Q9XGB8	Q9xgb8 triticum ae
36	36	58.1	860	10	Q9LXU3	Q9lxu3 arabidopsis
37	36	58.1	861	10	Q9LXU4	Q9lxu4 arabidopsis
38	36	58.1	865	10	Q9XGB2	Q9xgb2 triticum ae
39	36	58.1	986	5	Q9U159	Q9u159 caenorhabdi
40	36	58.1	1073	2	Q9L5V8	Q9l5v8 prevotella
41	36	58.1	1555	5	Q9U8S7	Q9u8s7 caenorhabdi
42	36	58.1	1874	5	P90749	P90749 caenorhabdi
43	36	58.1	1973	5	Q17465	Q17465 caenorhabdi
44	36	58.1	4809	2	Q93HH0	Q93hh0 streptomyce
45	35	56.5	186	9	Q94MX7	Q94mx7 haemophilus

ALIGNMENTS

RESULT 1
Q96HI7
ID Q96HI7 PRELIMINARY; PRT; 383 AA.
AC Q96HI7;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE SIMILAR TO CHITINASE 3-LIKE 1 (CARTILAGE GLYCOPROTEIN-39).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BRAIN, AND GLIOBLASTOMA WITH EGFR AMPLIFICATION;
RA Strausberg R.;
RL Submitted (MAY-2001) to EMBL/GenBank/DDBJ databases.
DR EMBL; BC008568; AAH08568.1; -.
SQ SEQUENCE 383 AA; 42625 MW; 6C9EF134BDC7C2D1 CRC64;

Query Match 100.0%; Score 62; DB 4; Length 383;
Best Local Similarity 100.0%; Pred. No. 0.0012;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RSFTLASSETGVG 13
||||||||||||| |
Db 263 RSFTLASSETGVG 275

RESULT 2
Q29411
ID Q29411 PRELIMINARY; PRT; 383 AA.
AC Q29411;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE 48 KDA HEPARIN-BINDING GLYCOPROTEIN.
OS Sus scrofa (Pig).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=SMOOTH MUSCLE;
RX MEDLINE=95286589; PubMed=7768982;
RA Shackelton L.M., Mann D.M., Millis A.J.;
RT "Identification of a 38-kDa heparin-binding glycoprotein (gp38k) in
RT differentiating vascular smooth muscle cells as a member of a group of
RT proteins associated with tissue remodeling.";
RL J. Biol. Chem. 270:13076-13083(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=SMOOTH MUSCLE;
RA Shackelton L.M., Mann D.M., Millis A.J.J.;
RL Submitted (JAN-1995) to the EMBL/GenBank/DDBJ databases.
DR EMBL; U17600; AAA86482.1; -.
DR EMBL; Z47803; CAA87764.1; -.
DR InterPro; IPR001223; Glyco_hydro_18.
DR Pfam; PF00704; Glyco_hydro_18; 1.
KW Heparin-binding.
SQ SEQUENCE 383 AA; 42443 MW; 3D1039E499108B0C4 CRC64;

Query Match 82.9%; Score 52; DB 6; Length 383;
Best Local Similarity 84.6%; Pred. No. 0.16;
Matches 11; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RSFTLASSEIGVG 13
:|||:|||: ||
Db 269 RSFILASSKIDVG 275

RESULT 3
O18949
ID O18949 PRELIMINARY; PRT; 332 AA.
AC O18949;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE CHITINASE-LIKE PROTEIN 1 (FRAGMENT).
GN CLP 1.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RA Recklies A.D., White C.;
RT "Expression of chitinase-like protein 1 (CLP-1) in bovine
RT chondrocytes.";
RL Submitted (JUN-1997) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AF011473; AAB64304.1; -.
DR InterPro; IPR001223; Glyco_hydro_18.
DR Pfam; PF00704; Glyco_hydro_18; 1.
FT NON_TER 1 1
FT NON_TER 332 332
SQ SEQUENCE 332 AA; 37434 MW; 03E1833248B4D0F CRC64;

Query Match 77.4%; Score 48; DB 6; Length 332;
Best Local Similarity 76.9%; Pred. No. 0.52;
Matches 10; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RSFTLASSEIGVG 13
:|:|:|:|:| ||
Db 239 RSYILASSKIDVG 251

RESULT
Q9WTV1

ID Q9WTV1 PRELIMINARY; PRT; 352 AA.
AC Q9WTV1;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE GLYCOPROTEIN 39 (FRAGMENT).
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciuromorpha; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LEWIS;
RA Wendling U., Boots A.M.H., van Eden W.;
RT "Cloning of the rat homologue of human cartilage glycoprotein-39 a
RT potential autoantigen in arthritis.";
RL Submitted (APR-2000) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AF062036; AAD22616.1; -.
DR HSSP; P37264; 1C1N.
DR InterPro; IPR000772; 2S-Globulin.
DR InterPro; IPR001223; Glyco_hydro_18.
DR Pfam; PF00704; Glyco_hydro_18; 1.
DR PRINTS; PR00551; 2SGLOBULIN.
FT NON_TER 1 1
FT NON_TER 352 352
SQ SEQUENCE 352 AA; 39391 MW; C80D9F6A0A936C CRC64;

Query Match 74.2%; Score 46; DB 11; Length 352;
Best Local Similarity 76.9%; Pred. No. 1.4;
Matches 10; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 RSFTLASSETGVG 13
:|||||||| ||
Db 242 KSFTLASSENGVG 254

RESULT 5
Q96F97
ID Q96F97 PRELIMINARY; PRT; 390 AA.
AC Q96F97;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE CHITINASE 3-LIKE 2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BRAIN, AND GLIOBLASTOMA WITH EGFR AMPLIFICATION.
RA Strausberg R.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DDBJ databases.
DR EMBL; BC011460; AAH11460.1; -.
SQ SEQUENCE 390 AA; 43529 MW; C801E4AF7A1A6CF CRC64;

Query Match 74.2%; Score 46; DB 4; Length 390;
Best Local Similarity 83.3%; Pred. No. 1.5;
Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 SFTLASSETGVG 13
:|||||:|| ||
Db 272 SFILASAETIVG 283

RESULT 6
Q9D7Q1
ID Q9D7Q1 PRELIMINARY; PRT; 390 AA.
AC Q9D7Q1;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)



```
DT   01-JUN-2001 (TrEMBLrel. 14, Created)
DE   ZYMOGEN-LIKE PROTEIN.
GN   ZYMOGEN.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6; TISSUE=BONE MARROW;
RX   MEDLINE=21385660; PubMed=11217851;
RA   Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA   Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA   Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA   Saito T., Kawai Y., Gojobori T., Hono H., Kasukawa T., Saito R.,
RA   Kadota K., Matsuda H., Ashburner M., Batalov S., Casavant T.,
RA   Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA   Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA   Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA   Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA   Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA   Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA   Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA   Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA   Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA   Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA   Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA   Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA   Hayashizaki Y.;
RT   "Functional annotation of a full-length mouse cDNA collection.";
RL   Nature 409:685-690(2001).
CC   -!- SIMILARITY: BELONGS TO CHITINASE CLASS II (FAMILY 18 OF GLYCOSYL
CC       HYDROLASES).
DR   EMBL; AK009012; BAB26025.1; -.
DR   MGD; MGI:1914546; Zymogen.
DR   InterPro; IPR001223; Glyco_hydro_18.
DR   InterPro; IPR001479; Chitinase_2.
DR   InterPro; IPR002283; Glyco_hydro_18.
DR   Pfam; PF00704; Glyco_hydro_18; 1.
DR   PRINTS; PR00051; 2SCLOBULIN.
DR   PROSITE; PS01095; CHITINASE_18; 1.
KW   Glycosidases; Hydrolase.
SQ   SEQUENCE   396 AA;  43413 MW;  82BB674B61AF7123 CRC64;


  Query Match             70.2%;  Score 43.5;  DB 11;  Length 396;
  Best Local Similarity   78.6%;  Pred. No. 4.7;
  Matches   11;  Conservative    1;  Mismatches    1;  Indels    1;  Gaps    1;

Qy       1 RSFTLA-SSEIGVG 13
           |||||| || :||
Db     269 RSFTLASSSLNGVG 282


RESULT   7
Q99J84
ID   Q99J84      PRELIMINARY;      PRT;   381 AA.
AC   Q99J84;
DT   01-JUN-2001 (TrEMBLrel. 14, Created)
DT   01-JUN-2001 (TrEMBLrel. 14, Last sequence update)
DT   01-OCT-2001 (TrEMBLrel. 15, Last annotation update)
DE   SIMILAR TO CHITINASE 3 LIKE 1 (CARTILAGE GLYCOPROTEIN 39).
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Strausberg R.;
RL   Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Strausberg R.;
RL   Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RA   Strausberg R.;
RL   Submitted (MAR-22-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC005611; AAH05611.1; -.
DR   EMBL; BC003780; AAH03780.1; -.
DR   EMBL; BC004734; AAH04734.1; -.
DR   InterPro; IPR001223; Glyco_hydro_18.
DR   Pfam; PF00704; Glyco_hydro_18; 1.
SQ   SEQUENCE   381 AA;  42979 MW;  EF6588C5AE9D4450 CRC64;


  Query Match             69.1%;  Score 43;  DB 11;  Length 381;
  Best Local Similarity   69.2%;  Pred. No. 5.6;
  Matches    9;  Conservative    2;  Mismatches    2;  Indels    1;  Gaps    1;

Qy       1 RSFTLASSETGVG 13
           |||||||| :|
Db     264 RSFTLASSENQLG 276


RESULT   8
Q9XED2
ID   Q9XED2      PRELIMINARY;      PRT;   807 AA.
AC   Q9XED2;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   STARCH BRANCHING ENZYME-1.
GN   SBE-1.
OS   Triticum aestivum (Wheat).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae;
OC   Triticeae; Triticum.
OX   NCBI_TaxID=4565;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=ENDOSPERM;
RA   Rahman S., Li Z., Abrahams S., Abbott D., Appels R., Morell M.K.;
RT   "Characterisation of a gene encoding wheat endosperm starch branching
RT   enzyme 1.";
RL   Theor. Appl. Genet. 98:156-163(1999).
DR   EMBL; AF076679; AAD30186.1; -.
DR   InterPro; IPR000461; Alpha-amylase.
DR   InterPro; IPR004193; Isoamylase_N.
DR   Pfam; PF00128; alpha-amylase; 1.
DR   Pfam; PF02922; Isoamylase_N; 1.
SQ   SEQUENCE   807 AA;  91390 MW;  B9E27F0158371813 CRC64;


  Query Match             66.1%;  Score 41;  DB 10;  Length 807;
  Best Local Similarity   53.8%;  Pred. No. 41;
  Matches    7;  Conservative    4;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 RSFTLASSETGVG 13
           | :|:|:: | |||
Db     43 RDYTMATAEDGVG 55


RESULT   9
Q9XED3
ID   Q9XED3      PRELIMINARY;      PRT;   829 AA.
AC   Q9XED3;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   STARCH BRANCHING ENZYME 1.
GN   SBE-1.
OS   Aegilops tauschii (Aegilops squarrosa).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae;
OC   Triticeae; Aegilops.
```

```
OX   NCBI_TaxID=37682;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Rahman S., Li Z., Abrahams S., Abbott D., Appels R., Morell M.K.;
RT   "Characterisation of a gene encoding wheat endosperm starch branching
RT   enzyme I.";
RL   Theor. Appl. Genet. 98:156-163(1999).
DR   EMBL; AF076980; AAD30187.1; -.
DR   InterPro; IPR000461; Alpha_amylase.
DR   InterPro; IPR004193; Isoamylase_N.
DR   Pfam; PF00128; alpha_amylase; 1.
DR   Pfam; PF02922; Isoamylase_N; 1.
SQ   SEQUENCE   829 AA;  94667 MW;  0E5534988F4877C7 CRC64;


  Query Match             66.1%;  Score 41;  DB 10;  Length 829;
  Best Local Similarity   58.8%;  Pred. No. 42;
  Matches    7;  Conservative    4;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 RSFTLASSEIGVG 13
           | :|:::: :||
Db     695 RDYIMAIAETGVG 708


RESULT  10
Q40074
ID   Q40074      PRELIMINARY;      PRT;   838 AA.
AC   Q40074;
DT   01-JUL-1997 (TrEMBLrel. 04, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   STARCH BRANCHING ENZYME I PRECURSOR (STARCH BRANCHING ENZYME I)
DE   (EC 2.4.1.18).
GN   SBE1 OR SBEID.
OS   Triticum aestivum (Wheat).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae;
OC   Triticeae; Triticum.
OX   NCBI_TaxID=4565;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CV. YIELDER;
RA   Repellin A., Nair K., Bana M., Chibbar R.N.;
RL   Plant Physiol. 114:1145-1145(1997).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CV. CHEYENNE; TISSUE=ENDOSPERM;
RA   McCue K.F., Workmaster W.L., Tanaka C.K., Anderson O.D.;
RT   "Starch Branching Enzymes Sbe1 and Sbe2 from Wheat (Triticum aestivum
RT   cv. Cheyenne): Molecular Characterization, Developmental Expression,
RT   and Homeolog Assignment by Differential PCR.";
RL   Submitted (JUL-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; Y12320; CAA72987.1; -.
DR   EMBL; AF286417; AAG27621.1; -.
DR   InterPro; IPR000461; Alpha_amylase.
DR   InterPro; IPR004193; Isoamylase_N.
DR   Pfam; PF00128; alpha_amylase; 1.
DR   Pfam; PF02922; Isoamylase_N; 1.
KW   Signal; Transferase; Glycosyltransferase.
FT   SIGNAL        1     61       POTENTIAL.
FT   CHAIN        62    838       POTENTIAL.
SQ   SEQUENCE   838 AA;  94577 MW;  85692153420FBA02E CRC64;


  Query Match             66.1%;  Score 41;  DB 10;  Length 838;
  Best Local Similarity   58.8%;  Pred. No. 42;
  Matches    7;  Conservative    4;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 RSFTLASSEIGVG 13
           | :|:::: :||
Db     695 RDYIMAIAETGVG 708


RESULT  11
Q9B4V8
ID   Q9B4V8      PRELIMINARY;      PRT;   867 AA.
AC   Q9B4V8;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   CHITOTRIOSIDASE PRECURSOR.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=96064695; PubMed=7592832;
RA   Boot R.G., Renkema G.H., Strijland A., van Zonneveld A.J., Aerts J.M.;
RT   "Cloning of a cDNA encoding chitotriosidase, a human chitinase
RT   produced by macrophages.";
RL   J. Biol. Chem. 270:26252-26256(1995).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=98421482; PubMed=9748265;
RA   Boot R.G., Renkema G.H., Verhoek M., Strijland A., Bliek J.,
RA   de Meulemeester T.M., Mannens M.M., Aerts J.M.;
RT   "The human chitotriosidase gene. Nature of inherited enzyme
RT   deficiency.";
RL   J. Biol. Chem. 273:25680-25685(1998).
CC   -!- SIMILARITY: BELONGS TO CHITINASE CLASS II (FAMILY 18.1 OF GLYCOSY
CC       HYDROLASES).
DR   EMBL; U02662; AAC50644.1; -.
DR   InterPro; IPR001579; Chitinase_2.
DR   InterPro; IPR001223; Glyco_hydro_18.
DR   Pfam; PF00704; Glyco_hydro_18; 1.
DR   PROSITE; PS01095; CHITINASE_18; 1.
KW   Glycosidase; Hydrolase; Signal.
FT   SIGNAL        1     21       POTENTIAL.
FT   CHAIN        22    387       CHITOTRIOSIDASE.
SQ   SEQUENCE   387 AA;  43133 MW;  83A272B0B5CE9D71 CRC64;


  Query Match             65.4%;  Score 41.5;  DB 4;  Length 387;
  Best Local Similarity   78.6%;  Pred. No. 17;
  Matches   11;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 RSFTLASSEIGVG 13
           |||||| ||:|| ||
Db     269 RSFTLASSSDIEVG 282


RESULT  12
Q13231
ID   Q13231      PRELIMINARY;      PRT;   466 AA.
AC   Q13231;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   CHITOTRIOSIDASE PRECURSOR.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=96064695; PubMed=7592832;
RA   Boot R.G., Renkema G.H., Strijland A., van Zonneveld A.J., Aerts J.M.;
RT   "Cloning of a cDNA encoding chitotriosidase, a human chitinase
RT   produced by macrophages.";
RL   J. Biol. Chem. 270:26252-26256(1995).
CC   -!- SIMILARITY: BELONGS TO CHITINASE CLASS II (FAMILY 18.1 OF GLYCOSY
CC       HYDROLASES).
DR   EMBL; U29615; AAC50246.1; -.
DR   InterPro; IPR001579; Chitinase_2.
DR   InterPro; IPR002557; Chitin_binding.
```

DR InterPro; IPR001223; Glyco_hydro_18.
DR Pfam; PF01607; Chitin_bind_3; 1.
DR Pfam; PF00704; Glyco_hydro_18; 1.
DR SMART; SM00494; ChtBD2; 1.
DR PROSITE; PS51910; CHITINASE_18; 1.
KW Glycosidase; Hydrolase; Signal.
FT SIGNAL 1 21 POTENTIAL.
FT CHAIN 22 466 CHITODEXTRINASE.
SQ SEQUENCE 466 AA; 51631 MW; 3A4120ED805E0A8A CRC64;

Query Match 65.3%; Score 40.5; DB 4; Length 466;
Best Local Similarity 78.6%; Pred. No. 21;
Matches 11; Conservative 1; Mismatches 1; Indels 1; Gaps 1;

Qy 1 RSPTLA-SSPTGVG 13
|||| | | |:| |
Db 269 RSPTLASSSTIRVG 282

RESULT 13
Q38137
ID Q38137 PRELIMINARY; PRT; 185 AA.
AC Q38137;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 16, Last annotation update)
DE ENDODEOXYRIBONUCLEASE.
OS Bacteriophage SP42.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Myoviridae.
OX NCBI_TaxID=10669;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=0;
RX MEDLINE=96041115; PubMed=7592802;
RA Goodrich-Blair H., Shub D.A.;
RT "The DNA polymerase genes of several HMU-bacteriophages have similar
RT group I introns with highly divergent open reading frames.";
RL Nucleic Acids Res. 23:3715-3723(1994).
DR EMBL; D04812; AAA50984.1; -.
DR InterPro; IPR000305; HNH.
DR InterPro; IPR003615; HNH_nuc.
DR InterPro; IPR004042; Intron_gpI.
DR Pfam; PF01844; HNH; 1.
DR SMART; SM00507; HNHc; 1.
DR SMART; SM00497; IENDI; 1.
SQ SEQUENCE 185 AA; 21299 MW; 5331A824B71EFDE9 CRC64;

Query Match 64.5%; Score 40; DB 9; Length 185;
Best Local Similarity 75.0%; Pred. No. 9.5;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 RSPTLASSEIGV 12
.| | ||| | |
Db 144 RSTIQASKEIGV 155

RESULT 14
Q9HPA0
ID Q9HPA0 PRELIMINARY; PRT; 261 AA.
AC Q9HPA0;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE VNG1735C.
GN VNG1735C;
OS Halobacterium sp. (strain NRC-1).
OC Archaea; Euryarchaeota; Halobacteriales; Halobacteriaceae;
OC Halobacterium.
OX NCBI_TaxID=64091;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20504483; PubMed=11016950;
RA Ng W.V., Kennedy S.P., Mahairas G.G., Berquist B., Pan M.,
RA Shukla H.D., Lasky S.R., Baliga N.S., Thorsson V., Sbrogna J.,
RA Swartzell S., Weir D., Hall J., Dahl T.A., Welti R., Goo Y.A.,
RA Leithauser B., Keller K., Cruz R., Danson M.J., Hough D.W.,
RA Maddocks D.G., Jablonski P.E., Krebs M.P., Angevine C.M., Dale H.,
RA Isenbarger T.A., Peck R.F., Pohlschroder M., Spudich J.L., Jung K.-H.,
RA Alam M., Freitas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D.,
RA Ebhardt H., Lowe T.M., Liang P., Riley M., Hood L., DasSarma S.;
RT "Genome sequence of Halobacterium species NRC-1.";
RL Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000).
DR EMBL; AE005079; AAG19970.1; -.
KW Complete proteome.
SQ SEQUENCE 261 AA; 28217 MW; 6F5304AB9F778789 CRC64;

Query Match 61.3%; Score 38; DB 17; Length 261;
Best Local Similarity 63.6%; Pred. No. 44;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 RSPTLASSEPG 11
|:||| |:||
Db 45 RTPTLVDSDPG 45

RESULT 15
O52483
ID O52483 PRELIMINARY; PRT; 441 AA.
AC O52483;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 16, Last annotation update)
DE PUTATIVE ATP BINDING PROTEIN.
GN WZ1.
OS Serratia marcescens.
OC Bacteria; Proteobacteria; Gamma subdivision; Enterobacteriaceae;
OC Serratia.
OX NCBI_TaxID=615;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=N28B;
RX MEDLINE=99172913; PubMed=10074083;
RA Sabti F., Climent N., Pique N., Sanchez C., Merino S., Rubires X.,
RA Aguilar A., Tomas J.M., Regue M.;
RT "Genetic analysis of the Serratia marcescens N28b O4 antigen gene
RT cluster.";
RL J. Bacteriol. 181:1883-1891(1999).
DR EMBL; AF038816; AAC00182.1; -.
DR InterPro; IPR003593; AAA.
DR InterPro; IPR003439; ABC_transporter.
DR InterPro; IPR001687; ATP_GTP_A.
DR Pfam; PF00005; ABC_tran; 1.
DR SMART; SM00382; AAA; 1.
KW ATP-binding.
SQ SEQUENCE 441 AA; 48197 MW; 56F7B4F1BDC2AB5D CRC64;

Query Match 61.3%; Score 38; DB 2; Length 441;
Best Local Similarity 88.9%; Pred. No. 60;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 SPTLASSEI 10
:|||||||:|
Db 331 NPTLASSEI 339

Search completed: August 26, 2002, 06:28:09
Job time: 1045 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 26, 2002, 05:23:29 ; Search time 46.22 Seconds
(without alignments)
6.870 Million cell updates/sec

Title: US-09-744-282-6
Perfect score: 62
Sequence: 1 RSFTLASSETGVG 13

Scoring table: BLOSUM62
Gap p 10.0 , Gapex 0.5

Searched: 231628 seqs, 24425694 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0
Maximum DB seq length: 2147000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/2/iaa/backfiles7.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	62	100.0	13	1	US-08-619-645-6	Sequence 6, Appli
2	62	100.0	13	2	US-08-634-493-6	Sequence 6, Appli
3	62	100.0	16	4	US-09-171-705-41	Sequence 41, Appl
4	46	74.2	385	2	US-08-694-915-2	Sequence 2, Appli
5	46	74.2	416	2	US-08-694-915-4	Sequence 4, Appli
6	43	69.4	9	1	US-08-619-645-2	Sequence 2, Appli
7	43	69.4	9	2	US-08-634-493-2	Sequence 2, Appli
8	43	69.4	125	1	US-08-478-039-69	Sequence 69, Appl
9	43	69.4	125	1	US-08-476-349A-69	Sequence 69, Appl
10	41	66.1	13	1	US-08-619-645-5	Sequence 5, Appli
11	41	66.1	13	2	US-08-634-493-5	Sequence 5, Appli
12	41	66.1	16	4	US-09-171-705-40	Sequence 40, Appl
13	40.5	65.3	373	4	US-09-039-198A-14	Sequence 14, Appl
14	40.5	65.3	373	4	US-09-039-198A-15	Sequence 15, Appl
15	40.5	65.3	387	2	US-08-486-839-6	Sequence 6, Appli
16	40.5	65.3	387	3	US-09-151-011-6	Sequence 6, Appli
17	40.5	65.3	387	4	US-09-443-623-6	Sequence 6, Appli
18	40.5	65.3	466	2	US-08-486-839-4	Sequence 4, Appli
19	40.5	65.3	466	3	US-09-151-011-4	Sequence 4, Appli
20	40.5	65.3	466	4	US-09-039-198A-2	Sequence 2, Appli
21	40.5	65.3	466	4	US-09-039-198A-4	Sequence 4, Appli
22	40.5	65.3	466	4	US-09-443-623-4	Sequence 4, Appli
23	38	61.3	16	4	US-09-171-705-42	Sequence 42, Appl
24	37	59.7	502	4	US-09-418-814-106	Sequence 106, App
25	36	58.1	82	2	US-08-778-251-22	Sequence 22, Appl
26	36	58.1	125	2	US-08-778-251-2	Sequence 2, Appli
27	35	56.5	686	2	US-08-849-480A-4	Sequence 4, Appli
28	34	54.8	268	2	US-08-860-577-4	Sequence 4, Appli
29	34	54.8	423	4	US-08-846-826A-2	Sequence 2, Appli
30	34	54.8	424	2	US-08-860-577-8	Sequence 8, Appli
31	34	54.8	471	3	US-09-043-627-10	Sequence 10, App
32	32	51.6	9	1	US-08-619-645-1	Sequence 1, Appli
33	32	51.6	9	2	US-08-634-493-1	Sequence 1, Appli
34	32	51.6	73	4	US-09-227-357-175	Sequence 175, App
35	32	51.6	240	2	US-08-114-555A-8	Sequence 8, Appli
36	32	51.6	240	3	US-08-559-397A-14	Sequence 14, Appl
37	32	51.6	297	2	US-08-580-545B-6	Sequence 6, Appli
38	32	51.6	297	4	US-09-262-654A-6	Sequence 6, Appli
39	32	51.6	563	1	US-08-106-761-4	Sequence 4, Appli
40	32	51.6	685	3	US-08-947-965-74	Sequence 74, Appl
41	32	51.6	831	2	US-09-047-026A-4	Sequence 4, Appli
42	32	51.6	976	3	US-08-560-005-2	Sequence 2, Appli
43	32	51.6	976	3	US-09-195-868-14	Sequence 14, Appl
44	32	51.6	976	4	US-09-418-540-2	Sequence 2, Appli
45	32	51.6	1098	4	US-08-946-994-17	Sequence 17, Appl

ALIGNMENTS

RESULT 1
US-08-619-645-6
; Sequence 6, Application US/08619645
; Patent No. 5736507
; GENERAL INFORMATION:
; APPLICANT: Boots, Anna M.H.
; APPLICANT: Verheijden, Gijsbertus F.M.
; TITLE OF INVENTION: N [illegible]-derivatives from
; TITLE OF INVENTION: [illegible] for use in immunotherapy of autoimmune
; TITLE OF INVENTION: disease
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Akzo Nobel Patent Department
; STREET: 1300 Piccard Drive, Suite 206
; CITY: Rockville
; STATE: Maryland
; COUNTRY: U.S.A.
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25(EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/619,645
; FILING DATE: 25-MAR-1996
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Gormley, Mary E.
; REGISTRATION NUMBER: 34,409
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 947-4433
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 13 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-619-645-6

Query Match 100.0%; Score 62; DB 1; Length 13;
Best Local Similarity 100.0%; Pred. No. 7.3e-06;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RSFTLASSETGVG 13
||||||||||||||
Db 1 RSFTLASSETGVG 13

```
RESULT   2
US-08-634-493-6
; Sequence 6, Application US/08634493
; Patent No. 5844449
;  GENERAL INFORMATION:
;    APPLICANT:  A.M.H. Boots
;    APPLICANT:  G.F.M. Verheijden
;    APPLICANT:  E.S. Bos
;    TITLE OF INVENTION:  No. 5844449e1 Peptides derived from autoantigen for use
;    TITLE OF INVENTION:  in Immunotherapy of Autoimmune Diseases
;    NUMBER OF SEQUENCES:  10
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Akzo No. 5844449e1 Patent Department
;      STREET:  1300 Piccard Drive, Suite 206
;      CITY:  Rockville
;      STATE:  Maryland
;      COUNTRY:  USA
;      ZIP:  20850
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/634,493
;      FILING DATE:  18-APR-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/619,645
;      FILING DATE:  25-MAR 1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/EP95/04203
;      FILING DATE:  25-OCT-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  NL 9420x1287
;      FILING DATE:  27-OCT 1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  NL 95200860
;      FILING DATE:  07-APR 1995
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (301) 258 5200
;      TELEFAX:  (301) 977 0847
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  13 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-634-493-6


  Query Match             100.0%;  Score 62;  DB 2;  Length 13;
  Best Local Similarity   100.0%;  Pred. No. 7.4e-06;
  Matches   13;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 RSPTLASSETGVG 13
           |||||||||||||
Db       1 RSPTLASSETGVG 13


RESULT   3
US-09-171-705-41
; Sequence 41, Application US/09171705
; Patent No. 6184204
; GENERAL INFORMATION:
;  APPLICANT: BOOTS, ANNA M.H.
;  APPLICANT: VERHEIJDEN, GILBERTUS F.M.
;  TITLE OF INVENTION: NOVEL PEPTIDES SUITABLE FOR USE IN ANTIGEN SPECIFIC
;  TITLE OF INVENTION: IMMUNOSUPPRESSIVE THERAPY
;  FILE REFERENCE: o/96198 US
;  CURRENT APPLICATION NUMBER: US/09/171,705
;  CURRENT FILING DATE: 1999-02-09
;  NUMBER OF SEQ ID NOS: 78
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 41
;   LENGTH: 16
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: DERIVED FROM
;   OTHER INFORMATION: SEQUENCE OF HUMAN CARTILAGE (HC)  30 FRAGMENT
US-09-171-705-41


  Query Match             100.0%;  Score 62;  DB 4;  Length 16;
  Best Local Similarity   100.0%;  Pred. No. 7.4e-06;
  Matches   13;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 RSPTLASSETGVG 13
           |||||||||||||
Db       2 RSPTLASSETGVG 14


RESULT   4
US-08-694-915-2
; Sequence 2, Application US/08694915
; Patent No. 5811535
; GENERAL INFORMATION:
;    APPLICANT:  Adamou, Julie
;    APPLICANT:  Kirkpatrick, Robert
;    APPLICANT:  Rosenberg, Martin
;    TITLE OF INVENTION:  HUMAN CARTILAGE GP39 LIKE GENE
;    NUMBER OF SEQUENCES:  7
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SmithKline Beecham Corporation
;      STREET:  709 Swedeland Road
;      CITY:  King of Prussia
;      STATE:  PA
;      COUNTRY:  USA
;      ZIP:  19406-2799
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ Version 1.5
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/694,915
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Ban, William I
;      REGISTRATION NUMBER:  34,844
;      REFERENCE/DOCKET NUMBER:  AIG50017
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  610-270-5219
;      TELEFAX:  610-270-5090
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  385 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    FRAGMENT TYPE:  N-terminal
;    ORIGINAL SOURCE:
US-08-694-915-2


  Query Match              74.2%;  Score 46;  DB 2;  Length 385;
```

Best Local Similarity 83.3%; Pred. No. 0.46;
Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 SFTLASSETVG 11
||||||:|| |
Db 267 SFTLASAETVG 278

RESULT 5
US-08-694-915-4
; Sequence 4, Application US/08694915
; Patent No. 5811538
; GENERAL INFORMATION:
; APPLICANT: Adamou, Julie
; APPLICANT: Kirkpatrick, Robert
; APPLICANT: Rosenberg, Martin
; TITLE OF INVENTION: HUMAN CARTILAGE GP39-LIKE GENE
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406-2799
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/694,915
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Han, William T
; REGISTRATION NUMBER: 34,544
; REFERENCE/DOCKET NUMBER: ATG50017
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-5219
; TELEFAX: 610 270-5090
; TELEX:
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 416 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
US-08-694-915-4

Query Match 74.2%; Score 46; DB 2; Length 416;
Best Local Similarity 83.3%; Pred. No. 0.5;
Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 SFTLASSETVG 11
||||||:|| |
Db 298 SFTLASAETVG 309

RESULT 6
US-08-619-645-2
; Sequence 2, Application US/08619645
; Patent No. 5736507
; GENERAL INFORMATION:
; APPLICANT: Boots, Anna M.H.
; APPLICANT: Verheijden, Gijsbertus F.M.
; TITLE OF INVENTION: No. 5736507el peptides derived from
; TITLE OF INVENTION: autoantigen for use in immunotherapy of autoimmune
; TITLE OF INVENTION: disease
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Akzo No. 5736507el Patent Department
; STREET: 1300 Piccard Drive, Suite 206
; CITY: Rockville
; STATE: Maryland
; COUNTRY: U.S.A.
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25(EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/619,645
; FILING DATE: 25-MAR-1996
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Gormley, Mary E.
; REGISTRATION NUMBER: 34,409
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 947-4433
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-619-645-2

Query Match 69.4%; Score 43; DB 1; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.7e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 FTLASSETG 11
|||||||||
Db 1 FTLASSETG 9

RESULT 7
US-08-634-493-2
; Sequence 2, Application US/08634493
; Patent No. 5843449
; GENERAL INFORMATION:
; APPLICANT: A.M.H. Boots
; APPLICANT: G.F.M. Verheijden
; APPLICANT: E.S. Bos
; TITLE OF INVENTION: No. 5843449el Peptides derived from autoantigen for use
; TITLE OF INVENTION: in Immunotherapy of Autoimmune Diseases
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Akzo No. 5843449el Patent Department
; STREET: 1300 Piccard Drive, Suite 206
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/634,493
; FILING DATE: 18-APR-1996
; PRIOR APPLICATION DATA:

```
;       APPLICATION NUMBER:  US 08/619,645
;       FILING DATE:  25-MAR-1996
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  PCT/EP95/04201
;       FILING DATE:  25-OCT-1995
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  NL 9420331287
;       FILING DATE:  27-OCT-1994
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  NL 952008860
;       FILING DATE:  07-APR-1995
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (301) 258-5200
;       TELEFAX:  (301) 977-0847
;   INFORMATION FOR SEQ ID NO:  2:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  9 amino acids
;       TYPE:  amino acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  peptide
US-08-634 494-2


  Query Match             69.4%;  Score 43;  DB 2;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 1.7e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3 FTLASSETG 11
           |||||||||
Db       1 FTLASSETG 9


RESULT   8
US-08-478-039-69
; Sequence 69, Application US/08478039
; Patent No. 5681722
;  GENERAL INFORMATION:
;    APPLICANT:  Newman, Roland A.
;    APPLICANT:  Hanna, Nabil
;    APPLICANT:  Kaab, Ronald W.
;    TITLE OF INVENTION:  Recombinant Antibodies for Human Therapy
;    NUMBER OF SEQUENCES:  114
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  BURNS, DOANE, SWECKER & MATHIS
;      STREET:  699 Prince St.
;      CITY:  Alexandria
;      STATE:  VA
;      COUNTRY:  USA
;      ZIP:  22313-1404
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/478,039
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/379,072
;      FILING DATE:  25-JAN-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/912,292
;      FILING DATE:  10-JUL-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/856,281
;      FILING DATE:  23-MAR-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/735,064
;      FILING DATE:  25-JUL-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Teskin Esq., Robin L.
;      REGISTRATION NUMBER:  35,030
;      REFERENCE/DOCKET NUMBER:  012712-160
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  703 836 6620
;      TELEFAX:  703-836-2021
;  INFORMATION FOR SEQ ID NO:  69:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  125 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  not relevant
;      TOPOLOGY:  not relevant
;    MOLECULE TYPE:  peptide
;    ORIGINAL SOURCE:
;      ORGANISM:  Monkey
;    POSITION IN GENOME:
;      CHROMOSOME/SEGMENT:  VH2 clone 2-10
US-08-478-039-69


  Query Match             69.4%;  Score 43;  DB 1;  Length 125;
  Best Local Similarity   72.7%;  Pred. No. 0.46;
  Matches    8;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3 FTLASSETGVG 13
           |:|::||||||
Db      27 FSLSTSETGVG 37


RESULT   9
US-08-476-349A-69
; Sequence 69, Application US/08476349A
; Patent No. 5750105
;  GENERAL INFORMATION:
;    APPLICANT:  Newman, Roland A.
;    APPLICANT:  Hanna, Nabil
;    APPLICANT:  Kaab, Ronald W.
;    TITLE OF INVENTION:  Recombinant Antibodies for Human Therapy
;    NUMBER OF SEQUENCES:  114
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  BURNS, DOANE, SWECKER & MATHIS
;      STREET:  699 Prince St.
;      CITY:  Alexandria
;      STATE:  VA
;      COUNTRY:  USA
;      ZIP:  22313-1404
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/476,349A
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/379,072
;      FILING DATE:  25-JAN-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/912,292
;      FILING DATE:  10-JUL-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/856,281
;      FILING DATE:  23-MAR-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/735,064
;      FILING DATE:  25-JUL-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Teskin Esq., Robin L.
;      REGISTRATION NUMBER:  35,030
;      REFERENCE/DOCKET NUMBER:  012712-161
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  703 836 6620
;      TELEFAX:  703-836-2021
```

```
;  INFORMATION FOR SEQ ID NO:  9:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  125 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  not relevant
;      TOPOLOGY:  not relevant
;    MOLECULE TYPE:  peptide
;    ORIGINAL SOURCE:
;      ORGANISM:  Monkey
;    POSITION IN GENOME:
;      CHROMOSOME/SEGMENT:  VH2 clone 2.10
US-08-476-349A-69


  Query Match             69.4%;  Score 43;  DB 1;  Length 125;
  Best Local Similarity   72.7%;  Pred. No. 0.46;
  Matches    8;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3 FTLASSEIGVS 13
           |:|::||||
Db      27 FSLSTSEIGVG 37


RESULT  10
US-08-619-645-5
; Sequence 5, Application US/08619645
; Patent No. 5736507
;  GENERAL INFORMATION:
;    APPLICANT:  Boots, Anna M.H.
;    APPLICANT:  Verheijden, Gijsbertus F.M.
;    TITLE OF INVENTION:  No. 5736507el peptides derived from
;    TITLE OF INVENTION:  autoantigen for use in immunotherapy of autoimmune
;    TITLE OF INVENTION:  disease
;    NUMBER OF SEQUENCES:  9
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Akzo No. 5736507el Patent Department
;      STREET:  1300 Piccard Drive, Suite 206
;      CITY:  Rockville
;      STATE:  Maryland
;      COUNTRY:  U.S.A.
;      ZIP:  20850
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25(EPO)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/619,645
;      FILING DATE:  25-MAR-1996
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION
;      NAME:  Formuler, Mary E.
;      REGISTRATION NUMBER:  34,429
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (301) 947-4433
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  13 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-619-645-5


  Query Match             66.1%;  Score 41;  DB 1;  Length 13;
  Best Local Similarity   100.0%;  Pred. No. 0.079;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 RSFTLASSE 9
           |||||||||
Db       5 RSFTLASSE 13


RESULT  11
US-08-634-493-5
; Sequence 5, Application US/08634493
; Patent No. 5843449
;  GENERAL INFORMATION:
;    APPLICANT:  A.M.H. Boots
;    APPLICANT:  G.F.M. Verheijden
;    APPLICANT:  E.S. Bos
;    TITLE OF INVENTION:  No. 5843449el Peptides derived from autoantigen for use
;    TITLE OF INVENTION:  in Immunotherapy of Autoimmune Diseases
;    NUMBER OF SEQUENCES:  10
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Akzo No. 5843449el Patent Department
;      STREET:  1300 Piccard Drive, Suite 206
;      CITY:  Rockville
;      STATE:  Maryland
;      COUNTRY:  USA
;      ZIP:  20850
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/634,493
;      FILING DATE:  18-APR-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/619,645
;      FILING DATE:  25-MAR-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  EP/EP95/04201
;      FILING DATE:  25-OCT-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  NL 9420231237
;      FILING DATE:  27-OCT-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  NL 9520008860
;      FILING DATE:  07-APR-1995
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (301) 258-5200
;      TELEFAX:  (301) 977-0847
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  13 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-634-493-5


  Query Match             66.1%;  Score 41;  DB 2;  Length 13;
  Best Local Similarity   100.0%;  Pred. No. 0.079;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 RSFTLASSE 9
           |||||||||
Db       5 RSFTLASSE 13


RESULT  12
US-09-171-705-40
; Sequence 40, Application US/09171705
; Patent No. 6184204
; GENERAL INFORMATION:
;  APPLICANT: BOOTS, ANNA M.H.
;  APPLICANT: VERHEIJDEN, GILBERTUS F.M.
;  TITLE OF INVENTION: NOVEL PEPTIDES SUITABLE FOR USE IN ANTIGEN SPECIFIC
;  TITLE OF INVENTION: IMMUNOSUPPRESSIVE THERAPY
;  FILE REFERENCE: 0/96198 US
;  CURRENT APPLICATION NUMBER: US/09/171,705
;  CURRENT FILING DATE: 1999-02-09
```

```
;  NUMBER OF SEQ ID NOS: 78
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 40
;    LENGTH: 16
;    TYPE: PRT
;    ORGANISM: Artificial Sequence
;    FEATURE:
;    OTHER INFORMATION: Description of Artificial Sequence: DERIVED FROM
;    OTHER INFORMATION: SEQUENCE OF HUMAN CARTILAGE (BC)  GP 18-CHITIN
US-09-171-705-40


  Query Match             66.1%;  Score 41;  DB 4;  Length 16;
  Best Local Similarity   100.0%;  Pred. No. 0.1;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 RSFTLASSE 9
           |||||||||
Db       8 RSFTLASSE 16


RESULT  13
US-09-039-198A-14
; Sequence 14, Application US/09039198A
; Patent No. 6200951
;  GENERAL INFORMATION:
;     APPLICANT:  Gray, Patrick W.
;     APPLICANT:  Tjoelker, Larry W.
;     TITLE OF INVENTION:  CHITINASE CHITIN BINDING FRAGMENTS
;     NUMBER OF SEQUENCES:  34
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray & Borun
;       STREET:  233 South Wacker Drive/6300 Sears Tower
;       CITY:  Chicago
;       STATE:  Illinois
;       COUNTRY:  United States of America
;       ZIP:  60606-6402
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.30
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/09/039,198A
;       FILING DATE:
;       CLASSIFICATION:  435
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Kin-Lauros, Li-Hsien
;       REGISTRATION NUMBER:  33,547
;       REFERENCE/DOCKET NUMBER:  2706/34391
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (312) 474-6300
;       TELEFAX:  (312) 474-0448
;   INFORMATION FOR SEQ ID NO:  14:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  473 amino acids
;       TYPE:  amino acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  peptide
US-09-039-198A-14


  Query Match             65.8%;  Score 40.5;  DB 4;  Length 473;
  Best Local Similarity   78.6%;  Pred. No. 5.1;
  Matches   11;  Conservative    1;  Mismatches    2;  Indels    1;  Gaps    1;

Qy       1 RSFTLA-SSETVG 13
           |||||  | | ||
Db     248 RSFTLASSSDTRVG 261


RESULT  14
```

```
US-09-039-198A-15
; Sequence 15, Application US/09039198A
; Patent No. 6200951
;  GENERAL INFORMATION:
;     APPLICANT:  Gray, Patrick W.
;     APPLICANT:  Tjoelker, Larry W.
;     TITLE OF INVENTION:  CHITINASE CHITIN BINDING FRAGMENTS
;     NUMBER OF SEQUENCES:  34
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray & Borun
;       STREET:  233 South Wacker Drive/6300 Sears Tower
;       CITY:  Chicago
;       STATE:  Illinois
;       COUNTRY:  United States of America
;       ZIP:  60606-6402
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.30
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/09/039,198A
;       FILING DATE:
;       CLASSIFICATION:  435
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Kin-Lauros, Li-Hsien
;       REGISTRATION NUMBER:  33,547
;       REFERENCE/DOCKET NUMBER:  2706/34391
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (312) 474-6300
;       TELEFAX:  (312) 474-0448
;   INFORMATION FOR SEQ ID NO:  15:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  473 amino acids
;       TYPE:  amino acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  peptide
US-09-039-198A-15


  Query Match             65.8%;  Score 40.5;  DB 4;  Length 473;
  Best Local Similarity   78.6%;  Pred. No. 5.1;
  Matches   11;  Conservative    1;  Mismatches    2;  Indels    1;  Gaps    1;

Qy       1 RSFTLA-SSETVG 13
           |||||  | | ||
Db     248 RSFTLASSSDTRVG 261


RESULT  15
US-08-486-859-6
; Sequence 6, Application US/08486859
; Patent No. 5928928
;  GENERAL INFORMATION:
;     APPLICANT:
;     TITLE OF INVENTION:  A human chitinase, its recombinant
;     TITLE OF INVENTION:  production, its use for decomposition of chitin, its use
;     TITLE OF INVENTION:  in therapy or prophylaxis against infections, and diseases
;     NUMBER OF SEQUENCES:  10
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Hoffmann & Baron
;       STREET:  459 Jericho Turnpike
;       CITY:  Jericho
;       STATE:  New York
;       COUNTRY:  United States of America
;       ZIP:  11753
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC DOS/MS DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.30
;       SOFTWARE:
```

```
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/486,839
;       FILING DATE:  07   June   1995
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Baron  Donald J.
;       REGISTRATION NUMBER:  [illegible]
;       REFERENCE/DOCKET NUMBER:  [illegible]
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (416) 822-[illegible]
;       TELEFAX:  (519) 822 3582
;   INFORMATION FOR SEQ ID NO:  6:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  387 amino acids
;       TYPE:  amino acid
;       STRANDEDNESS:  unknown
;       TOPOLOGY:  unknown
;     MOLECULE TYPE:  protein
;     HYPOTHETICAL:  NO
US/08/486,839-6


  Query Match              65.4%;  Score 40.5;  DB 2;  Length 387;
  Best Local Similarity    78.6%;  Pred. No. 5.8;
  Matches   11;  Conservative    1;  Mismatches    1;  Indels    1;  Gaps    1;

Qy       1 RSFTLA SSEDVC 13
             |||||| ||:  |:
Db     269 RSFTLASSSERVC 282


Search completed: August 26, 2002, 06:23:29
Job time: 870 sec
```

OC Nephropoidea; Nephropidae; Nephrops.
OX NCBI_TaxID=6829;
RN [1]
RP SEQUENCE.
RC TISSUE=Stomach;
RX MEDLINE=92089847; PubMed=1541966;
RA Barrett F., Keane T., Smithwick E., McGlynn E., Van Wramelcort A.;
RT "Structure and biological activity of crustacean gastroinhibitory
RT peptides identified with antibodies to gastrin/cholecystokinin.";
RL Biochimie 73:1249-1253(1991).
CC -!- FUNCTION: MAY CONTROL DIGESTION PROCESSES IN CRUSTACEANS.
CC -!- SIMILARITY: BELONGS TO THE GASTRIN/CCK FAMILY.
DR PIR; A49698; A49698.
KW Hormone.
SQ SEQUENCE 9 AA; 1098 MW; 0EC79CAB6D867B7B CRC64;

Query Match 44.4%; Score 15; DB 1; Length 9;
Best Local Similarity 50.0%; Pred. No. 1e-05;
Matches 4; Conservative 1; Mismatches 2; Indels 1; Gaps 0;

Qy 2 GRSFL 7
|:: |
Db 4 GQHFWL 9

RESULT 6
FAR5_CALVO
ID FAR5_CALVO STANDARD; PRT; 9 AA.
AC P41866;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE CalliFMRFamide 5.
OS Calliphora vomitoria (Blue blowfly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Oestroidea; Calliphoridae; Calliphora.
OX NCBI_TaxID=27454;
RN [1]
RP SEQUENCE.
RC TISSUE=Thoracic ganglion;
RX MEDLINE=92196111; PubMed=1549595;
RA Duve H., Johnsen A.H., Sewell J.C., Scott A.G., Orchard I.;
RA Rehfeld J.F., Thorpe A.;
RT "Isolation, structure, and activity of -Phe-Met-Arg-Phe-NH2
RT neuropeptides (designated calliFMRFamides) from the blowfly
RT Calliphora vomitoria.";
RL Proc. Natl. Acad. Sci. U.S.A. 89:2326-2330(1992).
CC -!- SIMILARITY: BELONGS TO THE FARP (FMRFAMIDE RELATED PEPTIDE)
CC FAMILY.
DR PIR; E41978; E41978.
KW Neuropeptide; Amidation.
FT MOD_RES 9 9 AMIDATION.
SQ SEQUENCE 9 AA; 1058 MW; 92D10699CAB6D867 CRC64;

Query Match 29.5%; Score 10; DB 1; Length 9;
Best Local Similarity 50.0%; Pred. No. 1e+02;
Matches 2; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 GRSF 5
|:: |
Db 4 GQDF 6

RESULT 7
FAR6_CALVO
ID FAR6_CALVO STANDARD; PRT; 9 AA.
AC P41867;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)

DE CalliFMRFamide 6.
OS Calliphora vomitoria (Blue blowfly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Oestroidea; Calliphoridae; Calliphora.
OX NCBI_TaxID=27454;
RN [1]
RP SEQUENCE.
RC TISSUE=Thoracic ganglion;
RX MEDLINE=92196111; PubMed=1549595;
RA Duve H., Johnsen A.H., Sewell J.C., Scott A.G., Orchard I.;
RA Rehfeld J.F., Thorpe A.;
RT "Isolation, structure, and activity of -Phe-Met-Arg-Phe-NH2
RT neuropeptides (designated calliFMRFamides) from the blowfly
RT Calliphora vomitoria.";
RL Proc. Natl. Acad. Sci. U.S.A. 89:2326-2330(1992).
CC -!- SIMILARITY: BELONGS TO THE FARP (FMRFAMIDE RELATED PEPTIDE)
CC FAMILY.
DR PIR; E41978; E41978.
KW Neuropeptide; Amidation.
FT MOD_RES 9 9 AMIDATION.
SQ SEQUENCE 9 AA; 1058 MW; 961D0699CAB6D865 CRC64;

Query Match 29.5%; Score 10; DB 1; Length 9;
Best Local Similarity 50.0%; Prod. No. 1e+02;
Matches 2; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 GRSF 5
|:: |
Db 4 GQDF 6

RESULT 8
FAR7_CALVO
ID FAR7_CALVO STANDARD; PRT; 9 AA.
AC P41862;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE CalliFMRFamide 7.
OS Calliphora vomitoria (Blue blowfly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Oestroidea; Calliphoridae; Calliphora.
OX NCBI_TaxID=27454;
RN [1]
RP SEQUENCE.
RC TISSUE=Thoracic ganglion;
RX MEDLINE=92196111; PubMed=1549595;
RA Duve H., Johnsen A.H., Sewell J.C., Scott A.G., Orchard I.;
RA Rehfeld J.F., Thorpe A.;
RT "Isolation, structure, and activity of -Phe-Met-Arg-Phe-NH2
RT neuropeptides (designated calliFMRFamides) from the blowfly
RT Calliphora vomitoria.";
RL Proc. Natl. Acad. Sci. U.S.A. 89:2326-2330(1992).
CC -!- SIMILARITY: BELONGS TO THE FARP (FMRFAMIDE RELATED PEPTIDE)
CC FAMILY.
DR PIR; G41978; G41978.
KW Neuropeptide; Amidation.
FT MOD_RES 9 9 AMIDATION.
SQ SEQUENCE 9 AA; 1087 MW; EBD10699CAB6D86A CRC64;

Query Match 29.5%; Score 10; DB 1; Length 9;
Best Local Similarity 50.0%; Prod. No. 1e+02;
Matches 2; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 GRSF 5
|:: |
Db 4 GQDF 6

```
RESULT  9
OCPI_OCTMI
ID   OCPI_OCTMI     STANDARD;      PRT;      4 AA.
AC   P58648;
DT   01-MAR-2002 (Rel. 41, Created)
DT   01-MAR-2002 (Rel. 41, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
DE   Cardioactive peptides Ocpi-1/Ocpi-2.
OS   Octopus minor (octopus).
OC   Eukaryota; Metazoa; Mollusca; Cephalopoda; Coleoidea; Octopoda;
OC   Incirrata; Octopodidae; Octopus.
OX   NCBI_TaxID=89776;
RN   [1]
RP   SEQUENCE, SYNTHESIS, MASS SPECTROMETRY, AND CHARACTERIZATION.
RC   TISSUE=Brain;
RX   PubMed=12876044;
RA   Iwakoshi E., Hisada M., Minakata H.;
RT   "Cardioactive peptides isolated from the brain of a Japanese octopus,
RT   Octopus minor.";
RL   Peptides 21:1625-1630(2000).
CC   -!- FUNCTION: cardioactive; has both positive chronotropic and
CC       inotropic effects on the heart. Ocpi-2 is a 1000-times less
CC       active than Ocpi-1.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- PTM: Ocpi-2 has D-Phe instead of L-Phe.
CC   -!- MASS SPECTROMETRY: MW=495.2; METHOD=MALDI.
KW   Hormone; D-amino acid.
FT   MOD_RES       2       2       D-PHENYLALANINE.
SQ   SEQUENCE   4 AA;  494 MW;  6AAB7C6B1B3B8B0B CRC64;


  Query Match          27.3%;  Score 12;  DB 1;  Length 4;
  Best Local Similarity  100.0%;  Pred. No. 1e+05;
  Matches    2;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 FG 2
           ||
Db       2 FG 3


RESULT  10
ALI4_CARMA
ID   ALI4_CARMA     STANDARD;      PRT;      5 AA.
AC   P81817;
DT   30-MAY 2000 (Rel. 39, Created)
DT   30-MAY 2000 (Rel. 39, Last sequence update)
DT   30-MAY 2000 (Rel. 39, Last annotation update)
DE   Carcinustatin 11.
OS   Carcinus maenas (Common shore crab) (Green crab).
OC   Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC   Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Brachyura;
OC   Eubrachyura; Portunoidea; Portunidae; Carcinus.
OX   NCBI_TaxID=6759;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Cerebral ganglion, and Thoracic ganglion;
RX   MEDLINE=98121193; PubMed=9461295;
RA   Duve H., Johnson A.H., Maestro J.-L., Scott A.G., Jaros P.P.,
RA   Thorpe A.;
RT   "Isolation and identification of multiple neuropeptides of the
RT   allatostatin superfamily in the shore crab Carcinus maenas.";
RL   Eur. J. Biochem. 250:727-734(1997).
CC   -!- FUNCTION: MAY ACT AS A NEUROTRANSMITTER OR NEUROMODULATOR.
CC   -!- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.
KW   Neuropeptide; Amidation; Multigene family.
FT   MOD_RES       5       5       AMIDATION (PHENYLAL).
SQ   SEQUENCE   5 AA;  596 MW;  62287C0FAB8B8B0B CRC64;


  Query Match          27.3%;  Score 12;  DB 1;  Length 5;
  Best Local Similarity  100.0%;  Pred. No. 1e+05;
  Matches    2;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 FG 2
           ||
Db       3 FG 4


RESULT  11
ALL2_CARMA
ID   ALL2_CARMA     STANDARD;      PRT;      7 AA.
AC   P81806;
DT   30-MAY 2000 (Rel. 39, Created)
DT   30-MAY 2000 (Rel. 39, Last sequence update)
DT   30-MAY 2000 (Rel. 39, Last annotation update)
DE   Carcinustatin 2.
OS   Carcinus maenas (Common shore crab) (Green crab).
OC   Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC   Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Brachyura;
OC   Eubrachyura; Portunoidea; Portunidae; Carcinus.
OX   NCBI_TaxID=6759;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Cerebral ganglion, and Thoracic ganglion;
RX   MEDLINE=98121193; PubMed=9461295;
RA   Duve H., Johnson A.H., Maestro J.-L., Scott A.G., Jaros P.P.,
RA   Thorpe A.;
RT   "Isolation and identification of multiple neuropeptides of the
RT   allatostatin superfamily in the shore crab Carcinus maenas.";
RL   Eur. J. Biochem. 250:727-734(1997).
CC   -!- FUNCTION: MAY ACT AS A NEUROTRANSMITTER OR NEUROMODULATOR.
CC   -!- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.
KW   Neuropeptide; Amidation; Multigene family.
FT   MOD_RES       7       7       AMIDATION (PHENYLAL).
SQ   SEQUENCE   7 AA;  779 MW;  67287C0FC0C0640620 CRC64;


  Query Match          27.3%;  Score 12;  DB 1;  Length 7;
  Best Local Similarity  100.0%;  Pred. No. 1e+05;
  Matches    2;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 FG 2
           ||
Db       5 FG 6


RESULT  12
ALL3_CARMA
ID   ALL3_CARMA     STANDARD;      PRT;      7 AA.
AC   P81806;
DT   30-MAY 2000 (Rel. 39, Created)
DT   30-MAY 2000 (Rel. 39, Last sequence update)
DT   30-MAY 2000 (Rel. 39, Last annotation update)
DE   Carcinustatin 3.
OS   Carcinus maenas (Common shore crab) (Green crab).
OC   Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC   Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Brachyura;
OC   Eubrachyura; Portunoidea; Portunidae; Carcinus.
OX   NCBI_TaxID=6759;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Cerebral ganglion, and thoracic ganglion;
RX   MEDLINE=98121193; PubMed=9461295;
RA   Duve H., Johnson A.H., Maestro J.-L., Scott A.G., Jaros P.P.,
RA   Thorpe A.;
RT   "Isolation and identification of multiple neuropeptides of the
RT   allatostatin superfamily in the shore crab Carcinus maenas.";
RL   Eur. J. Biochem. 250:727-734(1997).
CC   -!- FUNCTION: MAY ACT AS A NEUROTRANSMITTER OR NEUROMODULATOR.
CC   -!- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.
KW   Neuropeptide; Multigene family.
SQ   SEQUENCE   7 AA;  796 MW;  C287C0FC0C84Z0640Z0 CRC64;


  Query Match          27.3%;  Score 12;  DB 1;  Length 7;
  Best Local Similarity  100.0%;  Pred. No. 1e+05;
```

Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FG 2
||
Db 5 FG 6

RESULT 14
ALL4_CARMA
ID ALL4_CARMA STANDARD; PRT; 7 AA.
AC P82168;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Carcinustatin 4.
OS Carcinus maenas (Common shore crab) (Green crab).
OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Brachyura;
OC Eubrachyura; Portunoidea; Portunidae; Carcinus.
OX NCBI_TaxID=6759;
RN [1]
RP SEQUENCE.
RC TISSUE=Cerebral ganglion and thoracic ganglion;
RX MEDLINE=98121193; PubMed=9461296;
RA Dircksen H., Tensen A.B., Maestro J. L., Scott A.G., Furos P.P.,
RA Thorpe A.;
RT "Isolation and identification of multiple neuropeptides of the
RT allatostatin superfamily in the shore crab Carcinus maenas.";
RL Eur. J. Biochem. 250:727-744(1997).
CC -!- FUNCTION: MAY ACT AS A NEUROTRANSMITTER OR NEUROMODULATOR.
CC -!- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.
KW Neuropeptide; Multigene family.
SQ SEQUENCE 7 AA; 729 MW; 672979C0B479A0C0 CRC64;

Query Match 27.3%; Score 12; DB 1; Length 7;
Best Local Similarity 100.0%; Pred. No. 1e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FG 2
||
Db 5 FG 6

RESULT 15
ALL5_CARMA
ID ALL5_CARMA STANDARD; PRT; 7 AA.
AC P82168;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Carcinustatin 5.
OS Carcinus maenas (Common shore crab) (Green crab).
OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Brachyura;
OC Eubrachyura; Portunoidea; Portunidae; Carcinus.
OX NCBI_TaxID=6759;
RN [1]
RP SEQUENCE.
RC TISSUE=Cerebral ganglion and thoracic ganglion;
RX MEDLINE=98121193; PubMed=9461296;
RA Dircksen H., Johnsen A.H., Maestro J. L., Scott A.G., Furos P.P.,
RA Thorpe A.;
RT "Isolation and identification of multiple neuropeptides of the
RT allatostatin superfamily in the shore crab Carcinus maenas.";
RL Eur. J. Biochem. 250:727-744(1997).
CC -!- FUNCTION: MAY ACT AS A NEUROTRANSMITTER OR NEUROMODULATOR.
CC -!- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.
KW Neuropeptide; Amidation; Multigene family.
FT MOD_RES 7 7 AMIDATION.
SQ SEQUENCE 7 AA; 729 MW; 672979C0B4764200 CRC64;

Query Match 27.3%; Score 12; DB 1; Length 7;
Best Local Similarity 100.0%; Pred. No. 1e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FG 2
||
Db 5 FG 6

RESULT 15
ALL7_CYDPO
ID ALL7_CYDPO STANDARD; PRT; 7 AA.
AC P82158;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Cydiastatin 7.
OS Cydia pomonella (Codling moth).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
OC Tortricoidea; Tortricidae; Olethreutinae; Cydia.
OX NCBI_TaxID=82600;
RN [1]
RP SEQUENCE.
RC TISSUE=Larva;
RX MEDLINE=98054539; PubMed=9342829;
RA Duve H., Johnsen A.H., Maestro J. L., Scott A.G., Winstanley D.,
RA Davey M., East P.D., Thorpe A.;
RT "Lepidopteran peptides of the allatostatin superfamily.";
RL Peptides 18:1501-1509(1997).
CC -!- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.
KW Neuropeptide; Amidation.
FT MOD_RES 7 7 AMIDATION.
SQ SEQUENCE 7 AA; 873 MW; 672879C0BB669350 CRC64;

Query Match 27.3%; Score 12; DB 1; Length 7;
Best Local Similarity 100.0%; Pred. No. 1e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FG 2
||
Db 5 FG 6

Search completed: August 26, 2002, 06:40:42
Job time: 949 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

Mpi-beta - protein vs the usual sw model

Run on: August 26, 2002, 4:22:24 ; Search time 1.5 Seconds
(without alignments)
11.455 Million cell updates/sec

Title: US-09-744-282-1
Perfect score: 44
Sequence: 1 EKSEVLAS 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 56,222 seqs, 17,664,9 residues

Total number of hits satisfying chosen parameters: 148

Minimum DB seq length: 0
Maximum DB seq length: 9

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: SPTREMBL_19:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_viral:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	17	38.6	8	2	Q9S6D5	Q9S6D5 escherichia
2	17	38.6	9	10	Q9AXH8	Q9AXH8 mesembryant
3	16	36.4	8	8	P83858	P83858 pseudomonas
4	16	36.4	9	11	Q9QVH9	Q9QVH9 mus sp. sup
5	15	34.1	8	2	P92221	P92221 pseudomonas
6	15	34.1	8	8	P92432	P92432 psathyrosta
7	15	34.1	8	8	P92478	P92478 taphrina v
8	15	34.1	8	8	P93985	P93985 aeolops co
9	15	34.1	8	8	P92404	P92404 lophopyrum
10	15	34.1	8	8	P92426	P92426 pseudoroegn
11	15	34.1	8	8	P93973	P93973 elymopyrum
12	15	34.1	8	8	P93979	P93979 elymopyrum
13	15	34.1	8	8	P92388	P92388 leucaena p
14	15	34.1	8	8	P92429	P92429 peridiscus
15	15	34.1	8	8	P92201	P92201 heteranthel
16	15	34.1	8	8	P92227	P92227 crithopsis
17	15	34.1	8	8	P93963	P93963 psathyr sta
18	15	34.1	8	8	P93961	P93961 psathyr sta
19	15	34.1	8	8	P92215	P92215 amblypyrum
20	15	34.1	8	8	P93981	P93981 crithopsis
21	15	34.1	8	8	P92431	P92431 aeolops ha
22	15	34.1	8	8	P92222	P92222 bromus tec
23	15	34.1	8	8	P92443	P92443 lactiflora
24	15	34.1	8	8	P92482	P92482 hordeum bra
25	15	34.1	8	8	P92484	P92484 hordeum mar
26	15	34.1	8	8	P92494	P92494 hordeum cri
27	15	34.1	8	8	P93965	P93965 secale stri
28	15	34.1	8	8	P93966	P93966 aegil ps sp
29	15	34.1	8	8	P93992	P93992 austra. pyr
30	15	34.1	8	8	P92441	P92441 elytrigia
31	15	34.1	8	8	P93955	P93955 festucopsis
32	15	34.1	8	8	P92211	P92211 agropyron c
33	15	34.1	8	8	P93959	P93959 hordeum ere
34	15	34.1	8	8	P93957	P93957 festucopsis
35	15	34.1	9	2	Q9K4M6	Q9K4M6 staphy. aur
36	14	31.8	9	4	Q16220	Q16220 homo sapien
37	14	31.8	9	15	Q85710	Q85710 toxo satiu
38	13	29.5	7	2	O07454	O07454 synech cocc
39	13	29.5	8	8	P92219	P92219 anistrapyr
40	13	29.5	8	8	Q94VC1	Q94VC1 catanus 2.1
41	13	29.5	9	2	Q93LE4	Q93LE4 helicobac.
42	13	29.5	9	2	Q93SC0	Q93SC0 salmonella
43	13	29.5	9	8	Q94VI8	Q94VI8 catanus ore
44	13	29.5	9	8	Q94VH4	Q94VH4 catanus ala
45	13	29.5	9	12	Q66545	Q66545 human herpe

ALIGNMENTS

RESULT 1
Q9S6D5
ID Q9S6D5 PRELIMINARY; PRT; 8 AA.
AC Q9S6D5;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE PUTATIVE IS30 TRANSPOSASE (FRAGMENT).
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN A2958;
RX MEDLINE=99194747; PubMed=10094716;
RA Rahn A., Drummelsmith J., Whittield C.;
RT "Conserved organization in the cps gene clusters for expression-1
RT Escherichia coli group 1 K antigens: relationship to the colanic acid
RT biosynthesis locus and the cps genes from Klebsiella pneumoniae.";
RL J. Bacteriol. 181:2307-2314(1999).
DR EMBL; AF118251; AAD30008.1; -.
FT NON_TER 8 8
SQ SEQUENCE 8 AA; 1011 MW; F21DC1A9D1841406 CRC64;

Query Match 38.6%; Score 17; DB 2; Length 8;
Best Local Similarity 75.0%; Pred. No. 5.6e-05;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 3 KSEV 6
|:||
Db 3 KTEV 6

RESULT 2
Q9AXH8
ID Q9AXH8 PRELIMINARY; PRT; 9 AA.
AC Q9AXH8;

```
ID   01 JUN , 01 (IrEMBLrel. 17, Created)
DT   01 JUN 2001 (IrEMBLrel. 17, Last sequence update)
DT   01 JUN 2001 (IrEMBLrel. 17, Last annotation update)
DE   IIP AQUAPORIN (FRAGMENT).
GN   MIPK.
OS   Mesembryanthemum crystallinum (Common ice plant).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC   Caryophyllidae; Caryophyllales; Aizoaceae; Mesembryanthemum.
OX   NCBI TaxID=9544;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Bennett J.M., Quigley F.L., Bohnert H.J.;
RT   "The promoter of mipK gene which encodes for a TIP aquaporin from
RT   Mesembryanthemum crystallinum.";
RL   Submitted (DEC 2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF306797; AAG49697.1; -.
KW   Porin.
FT   NON_TER       9       9
SQ   SEQUENCE   9 AA;  909 MW;  C3ED0B79CED728ZD CRC64;


  Query Match             88.9%;  Score 17;  DB 1;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 5.6e+04;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 PGE 3
           ||
Db       2 PGE 4


RESULT   3
P82858
ID   P82858       PRELIMINARY;      PRT;     8 AA.
AC   P82858;
DT   01 MAR 2001 (IrEMBLrel. 16, Created)
DT   01 MAR 2001 (IrEMBLrel. 16, Last sequence update)
DT   01 JUN 2001 (IrEMBLrel. 17, Last annotation update)
DE   URICASE (EC 1.7.3.3) (URATE OXIDASE) (FRAGMENT).
OS   Pycnidia recondita L. sp. trisoti.
OC   Eukaryota; Fungi; Basidiomycota; Urediniomycetes; Uredinomycetidae;
OC   Uredinales; Pucciniaceae; Puccinia.
OX   NCBI TaxID=142079;
RN   [1]
RP   SEQUENCE, FUNCTION, AND DEVELOPMENTAL STAGE.
RC   TISSUE=SPORE;
RA   Aguilar M., Mollaghini P., Pineda M.;
RL   Submitted (NOV 2000) to the SWISS-PROT data bank.
CC   !  FUNCTION: INVOLVED IN A CT PARASITE RELATIONSHIP BETWEEN PLANTS
CC      AND FUNGI.
CC   !  CATALYTIC ACTIVITY: URATE + O(2) + H(2)O = 5-HYDROXYISOURATE
CC      H(2)O(2) (5-HYDROXYISOURATE HYDROLYSE TO FORM ALLANTOIN).
CC   !  SUBCELLULAR LOCATION: PEROXISOMAL (BY SIMILARITY).
CC   !  DEVELOPMENTAL STAGE: GERMINATION.
CC   !  SIMILARITY: BELONGS TO THE URICASE FAMILY.
DR   InterPro; IPR002042; Uricase.
DR   PROSITE; PS00366; URICASE; PARTIAL.
KW   Oxidoreductase; Purine metabolism; Peroxisome.
FT   NON_TER       8       8
SQ   SEQUENCE   8 AA;  777 MW;  98C1AD0735B9G761 CRC64;


  Query Match             88.9%;  Score 16;  DB 1;  Length 8;
  Best Local Similarity   100.0%;  Pred. No. 5.6e+04;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       5 FIAS 9
             ||||
Db       4 FIAS 7


RESULT   4
Q9QVH9
ID   Q9QVH9       PRELIMINARY;      PRT;     9 AA.
AC   Q9QVH9;
DT   01 MAY 2000 (IrEMBLrel. 1x, Created)
DT   01 MAY 2000 (IrEMBLrel. 1x, Last sequence update)
DT   01 MAY 2000 (IrEMBLrel. 1x, Last annotation update)
DE   SUPERINDUCIBLE PROTEIN 24, SSP24 (CYCLOPHILIN 8-MCI 3 (FRAG
OS   Mus sp.
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata, Euteleost
OC   Mammalia; Eutheria; Rodentia; Sciuromorpha, Muridae; Murinae; M
OX   NCBI TaxID=1094;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE 92041730; PubMed 1432154;
RA   Davis T.R., Tabaraban C., Bruns K., Ramirez R.L., Nelson Res. ; M.
RT   "Basic fibroblast growth factor induces 82 fibroblast-like cells
RT   and secrete a cyclophilin-like protein and beta-2 MCP-1 from
RL   Biochim. Biophys. Acta 1096:143-152(1991).
SQ   SEQUENCE   9 AA;  964 MW;  1C2A223A4D0E7C0 CRC64;


  Query Match             88.9%;  Score 16;  DB 1;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 5.6e+04;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       5 FIAS 9
             ||||
Db       5 FIGN 9


RESULT   5
P72221
ID   P72221       PRELIMINARY;      PRT;     8 AA.
AC   P72221;
DT   01 FEB 1997 (IrEMBLrel. 02, Created)
DT   01 FEB 1997 (IrEMBLrel. 02, Last sequence update)
DT   01 DEC 2001 (IrEMBLrel. 20, Last annotation update)
DE   ALGINATE LYASE (FRAGMENT).
GN   ALY.
OS   Pseudomonas sp. (strain 'S ALG 9).
OC   Bacteria; Proteobacteria.
OX   NCBI TaxID=80208;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=S ALG 9;
RA   Fujiyama K.;
RL   Submitted (OCT 1994) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=S-ALG 9;
RX   MEDLINE 93329356; PubMed 8393113;
RA   Maki H., Mori A., Fujiyama K., Kinoshita S., Yoshida T.;
RT   "Structural sequence analysis and expression in Escherichia coli
RT   gene encoding an alginate lyase from Pseudomonas sp. S ALG 9.";
RL   J. Gen. Microbiol. 139:987-993(1993).
DR   EMBL; D38469; BAA21704.1; -.
KW   Lyase.
FT   NON_TER       1       1
FT   NON_TER       8       8
SQ   SEQUENCE   8 AA;  841 MW;  4F0134-7A5704DB CRC64;


  Query Match             44.4%;  Score 8;  DB 1;  Length 8;
  Best Local Similarity   44.4%;  Pred. No. 5.6e+04;
  Matches    4;  Conservative    4;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       5 FIAS 9
             ::::
Db       1 YIHS 5


RESULT   6
P92422
ID   P92422       PRELIMINARY;      PRT;     8 AA.
```

```
AC   P92422; P92420
DT   01-MAY-1997 (TrEMBLrel. 03, Created)
DT   01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT   01-JAN-2000 (TrEMBLrel. 05, Last annotation update)
DE   CHLOROPLAST RUBA GENE (FRAGMENT).
GN   PEID.
OS   Psathyrostachys fragilis.
OG   Chloroplast.
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae;
OC   Triticeae; Psathyrostachys.
OX   NCBI_TaxID=4572;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN B4572, AND B672; TISSUE=LEAVES;
RA   Peterson G., Soreng R.;
RL   Submitted (JAN-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; Z77754; CAB01308.1; -.
DR   EMBL; Z77745; CAB01305.1; -.
KW   Chloroplast.
FT   NON_TER       1       1
SQ   SEQUENCE   8 AA;  878 MW;  1E572877241A745B3 CRC64;


  Query Match             34.1%;  Score 15;  DB 8;  Length 8;
  Best Local Similarity   60.0%;  Pred. No. 5.6e+05;
  Matches    3;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       5 RSFTL 7
             : | ||
Db       1 KSLIL 5


RESULT   7
P92423
ID   P92423       PRELIMINARY;      PRT;     8 AA.
AC   P92423;
DT   01-MAY-1997 (TrEMBLrel. 03, Created)
DT   01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   PEID PROTEIN (FRAGMENT).
GN   PEID.
OS   Baynaldia villosa.
OG   Chloroplast.
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PAC clade;
OC   Pooideae; Anthropogoneae; Baynaldia.
OX   NCBI_TaxID=40247;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN B5561; TISSUE=LEAF;
RX   MEDLINE=97271648; PubMed=9126564;
RA   Peterson G., Soreng R.;
RT   "Phylogenetic analysis of the Triticeae (Poaceae) based on rpoA
RT   sequence data.";
RL   Mol. Phylogenet. Evol. 7:217-230(1997).
DR   EMBL; Z77741; CAB01302.1; -.
KW   Chloroplast.
FT   NON_TER       1       1
SQ   SEQUENCE   8 AA;  878 MW;  1E572877241A745B3 CRC64;


  Query Match             34.1%;  Score 15;  DB 8;  Length 8;
  Best Local Similarity   60.0%;  Pred. No. 5.6e+05;
  Matches    3;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       5 RSFTL 7
             : | ||
Db       1 KSLIL 5


RESULT   8
P92405
ID   P92405       PRELIMINARY;      PRT;     8 AA.
AC   P92405;
DT   01-MAY-1997 (TrEMBLrel. 03, Created)
DT   01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   PEID PROTEIN (FRAGMENT).
GN   PEID.
OS   Aegilops comosa.
OG   Chloroplast.
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae;
OC   Triticeae; Aegilops.
OX   NCBI_TaxID=4485;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN B6623; TISSUE=LEAF;
RX   MEDLINE=97271648; PubMed=9126564;
RA   Peterson G., Soreng R.;
RT   "Phylogenetic analysis of the Triticeae (Poaceae) based on rpoA
RT   sequence data.";
RL   Mol. Phylogenet. Evol. 7:217-230(1997).
DR   EMBL; Z77742; CAB01305.1; -.
KW   Chloroplast.
FT   NON_TER       1       1
SQ   SEQUENCE   8 AA;  878 MW;  1E572877241A745B3 CRC64;


  Query Match             34.1%;  Score 15;  DB 8;  Length 8;
  Best Local Similarity   60.0%;  Pred. No. 5.6e+05;
  Matches    3;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       5 RSFTL 7
             : | ||
Db       1 KSLIL 5


RESULT   9
P92404
ID   P92404       PRELIMINARY;      PRT;     8 AA.
AC   P92404;
DT   01-MAY-1997 (TrEMBLrel. 03, Created)
DT   01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   PEID PROTEIN (FRAGMENT).
GN   PEID.
OS   Lophopyrum elongatum (Tall wheatgrass) (Agropyron elongatum).
OG   Chloroplast.
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae;
OC   Triticeae; Lophopyrum.
OX   NCBI_TaxID=4588;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN B6692; TISSUE=LEAF;
RX   MEDLINE=97271648; PubMed=9126564;
RA   Peterson G., Soreng R.;
RT   "Phylogenetic analysis of the Triticeae (Poaceae) based on rpoA
RT   sequence data.";
RL   Mol. Phylogenet. Evol. 7:217-230(1997).
DR   EMBL; Z77734; CAB01308.1; -.
KW   Chloroplast.
FT   NON_TER       1       1
SQ   SEQUENCE   8 AA;  878 MW;  1E572877241A745B3 CRC64;


  Query Match             34.1%;  Score 15;  DB 8;  Length 8;
  Best Local Similarity   60.0%;  Pred. No. 5.6e+05;
  Matches    3;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       5 RSFTL 7
             : | ||
Db       1 KSLIL 5
```

```
RESULT  10
P02426
ID   P02426       PRELIMINARY;      PRT:     8 AA.
AC   P02426;
DT   01-MAY-1997 (Rel. 03, Created)
DT   01-MAY-1997 (Rel. 03, Last sequence update)
DT   01-DEC-2001 (Rel. 19, Last annotation update)
DE   PETD PROTEIN (FRAGMENT).
GN   PETD.
OS   Pseudotrossomeria sprucata.
OC   Chloroplast.
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae;
OC   Triticeae; Pseudotrossomeria.
OX   NCBI_TaxID=5664;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=P0082; TISSUE=LEAF;
RX   MEDLINE=97271648; PubMed=9126564;
RA   Peterson G., Sebera G.;
RT   "Phylogenetic analysis of the Triticeae (Poaceae) based on rpoA
RT   sequence data.";
RL   Mol. Phylogenet. Evol. 7:217-230(1997).
DR   EMBL; Z77744; CAB01411.1; -.
KW   Chloroplast.
FT   NON_TER       1       1
SQ   SEQUENCE   8 AA;  878 MW;  1E372877A2758B4 CRC64;


  Query Match             44.4%;  Score 16;  DB 8;  Length 8;
  Best Local Similarity   66.7%;  Pred. No. 5.6e+05;
  Matches    8;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       3 RSFTL 7
           :| ||
Db       1 KSILL 5


RESULT  11
P93974
ID   P93974       PRELIMINARY;      PRT:     8 AA.
AC   P93974;
DT   01-MAY-1997 (Rel. 03, Created)
DT   01-MAY-1997 (Rel. 03, Last sequence update)
DT   01-DEC-2001 (Rel. 19, Last annotation update)
DE   PETD PROTEIN (FRAGMENT).
GN   PETD.
OS   Eromopyrum distans.
OC   Chloroplast.
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae;
OC   Triticeae; Eromopyrum.
OX   NCBI_TaxID=58666;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=B5552; TISSUE=LEAF;
RX   MEDLINE=97271648; PubMed=9126564;
RA   Peterson G., Sebera G.;
RT   "Phylogenetic analysis of the Triticeae (Poaceae) based on rpoA
RT   sequence data.";
RL   Mol. Phylogenet. Evol. 7:217-230(1997).
DR   EMBL; Z77745; CAB01414.1; -.
KW   Chloroplast.
FT   NON_TER       1       1
SQ   SEQUENCE   8 AA;  878 MW;  1E372877A2758B4 CRC64;


  Query Match             44.4%;  Score 16;  DB 8;  Length 8;
  Best Local Similarity   66.7%;  Pred. No. 5.6e+05;
  Matches    8;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       3 RSFTL 7
           :| ||
Db       1 KSILL 5


RESULT  12
P05970
ID   P05970       PRELIMINARY;      PRT:     8 AA.
AC   P05970;
DT   01-MAY-1997 (Rel. 03, Created)
DT   01-MAY-1997 (Rel. 03, Last sequence update)
DT   01-DEC-2001 (Rel. 19, Last annotation update)
DE   PETD PROTEIN (FRAGMENT).
GN   PETD.
OS   Eromopyrum triticeum.
OC   Chloroplast.
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae;
OC   Triticeae; Eromopyrum.
OX   NCBI_TaxID=58665;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=B5554; TISSUE=LEAF;
RX   MEDLINE=97271648; PubMed=9126564;
RA   Peterson G., Sebera G.;
RT   "Phylogenetic analysis of the Triticeae (Poaceae) based on rpoA
RT   sequence data.";
RL   Mol. Phylogenet. Evol. 7:217-230(1997).
DR   EMBL; Z77746; CAB01415.1; -.
KW   Chloroplast.
FT   NON_TER       1       1
SQ   SEQUENCE   8 AA;  878 MW;  1E372877A2758B4 CRC64;


  Query Match             44.4%;  Score 16;  DB 8;  Length 8;
  Best Local Similarity   66.7%;  Pred. No. 5.6e+05;
  Matches    8;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       3 RSFTL 7
           :| ||
Db       1 KSILL 5


RESULT  13
P92388
ID   P92388       PRELIMINARY;      PRT:     8 AA.
AC   P92388;
DT   01-MAY-1997 (Rel. 03, Created)
DT   01-MAY-1997 (Rel. 03, Last sequence update)
DT   01-DEC-2001 (Rel. 19, Last annotation update)
DE   PETD PROTEIN (FRAGMENT).
GN   PETD.
OS   Bernardia persica.
OC   Chloroplast.
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae;
OC   Triticeae; Bernardia.
OX   NCBI_TaxID=37678;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=B5556; TISSUE=LEAF;
RX   MEDLINE=97271648; PubMed=9126564;
RA   Peterson G., Sebera G.;
RT   "Phylogenetic analysis of the Triticeae (Poaceae) based on rpoA
RT   sequence data.";
RL   Mol. Phylogenet. Evol. 7:217-230(1997).
DR   EMBL; Z77748; CAB01423.1; -.
KW   Chloroplast.
FT   NON_TER       1       1
SQ   SEQUENCE   8 AA;  878 MW;  1E372877A2758B4 CRC64;


  Query Match             44.4%;  Score 16;  DB 8;  Length 8;
  Best Local Similarity   66.7%;  Pred. No. 5.6e+05;
```

```
  Matches   7;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy        6 RSFIL 7
            :| ||
Db        1 KSLIL 5


RESULT  14
P92428
ID   P92428      PRELIMINARY;      PRT;     8 AA.
AC   P92428;
DT   01-MAY-1997 (rEMBLrel. 04, Created)
DT   01-MAY-1997 (rEMBLrel. 04, Last sequence update)
DT   01-DEC-2001 (rEMBLrel. 19, Last annotation update)
DE   FEID PROTEIN (FRAGMENT).
GN   PEID.
OS   Peridiothyon sanctum.
OC   Chloroplast.
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae;
OC   Triticeae; Peridiothyon.
OX   NCBI_TaxID=3763;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Sanct; TISSUE=LEAF;
RX   MEDLINE=97271149; PubMed=9126564;
RA   Peterson L., Roberts C.;
RT   "Phylogenetic analysis of the Triticeae (Poaceae) based on rpoA
RT   sequence data.";
RL   Mol. Phylogenet. Evol. 7:25-32 (1997).
DR   EMBL; Z77749; CAA81526.1; -.
KW   Chloroplast.
FT   NON_TER                1
SQ   SEQUENCE   8 AA;  878 MW;  1E72877C1A75B3C0 CRC64;


  Query Match             84.1%;  Score 15;  DB 8;  Length 8;
  Best Local Similarity   60.0%;  Prod. No. 5.6e-05;
  Matches    4;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy        6 RSFIL 7
            :  |
Db        1 KSLIL 5


RESULT  15
P92491
ID   P92491      PRELIMINARY;      PRT;     8 AA.
AC   P92491;
DT   01-MAY-1997 (rEMBLrel. 04, Created)
DT   01-MAY-1997 (rEMBLrel. 04, Last sequence update)
DT   01-DEC-2001 (rEMBLrel. 19, Last annotation update)
DE   FEID PROTEIN (FRAGMENT).
GN   PEID.
OS   Heterantherium oligerum.
OC   Chloroplast.
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae;
OC   Triticeae; Heterantherium.
OX   NCBI_TaxID=3763;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Oligerum; TISSUE=LEAF;
RX   MEDLINE=97271611; PubMed=9126564;
RA   Peterson L., Roberts C.;
RT   "Phylogenetic analysis of the Triticeae (Poaceae) based on rpoA
RT   sequence data.";
RL   Mol. Phylogenet. Evol. 7:25-32 (1997).
DR   EMBL; Z77750; CAA81529.1; -.
KW   Chloroplast.
FT   NON_TER       1       1
SQ   SEQUENCE   8 AA;  878 MW;  1E72877C1A75B3C0 CRC64;


  Query Match             84.1%;  Score 15;  DB 8;  Length 8;
  Best Local Similarity   60.0%;  Prod. No. 5.6e-05;
  Matches    4;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy        6 RSFIL 7
            :| ||
Db        1 KSLPL 5


Search completed: August 26, 2002, 06:39:44
Job time: 975 sec
```

Geac to version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 26, 2002, 06:22:19 ; Search time 511.08 Seconds
(without alignments)
8.947 Million cell updates/sec

Title: US-09-744-282-6
Perfect score: 62
Sequence: 1 RSPTLASSELSVG 13

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 960,193 seqs, 351085561 residues

Total number of hits satisfying chosen parameters: 960193

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Patents_AA_Main.*
1: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*
15: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep:*
16: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep:*
17: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep:*
18: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep:*
19: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep:*
20: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep:*
21: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep:*
22: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep:*
23: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep:*
24: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*
25: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep:*
26: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	62	100.0	13	14	US-09-985-898-1	Sequence 1, Appli
2	62	100.0	13	15	US-09-158-900-6	Sequence 6, Appli
3	62	100.0	13	18	US-09-405-745-1	Sequence 1, Appli
4	62	100.0	13	21	US-09-744-170-3	Sequence 3, Appli
5	62	100.0	13	21	US-09-744-282-6	Sequence 6, Appli
6	62	100.0	13	23	US-09-985-065-1	Sequence 1, Appli
7	62	100.0	13	14	US-08-985-898-2	Sequence 2, Appli
8	62	100.0	15	18	US-09-405-745-2	Sequence 2, Appli
9	62	100.0	15	23	US-09-985-065-2	Sequence 2, Appli
10	62	100.0	16	13	US-08-985-898-3	Sequence 3, Appli
11	62	100.0	16	18	US-09-405-745-3	Sequence 3, Appli
12	62	100.0	16	23	US-09-985-065-3	Sequence 3, Appli
13	62	100.0	44	26	US-60-160-189-9814	Sequence 9814, Ap
14	62	100.0	44	26	US-60-160-203-5456	Sequence 5456, Ap
15	62	100.0	44	26	US-60-169-840-6915	Sequence 6915, Ap
16	62	100.0	44	26	US-60-169-867-6065	Sequence 6065, Ap
17	62	100.0	56	1	PCT-US01-14827-14284	Sequence 14284, A
18	62	100.0	199	1	PCT-US01-09616-56	Sequence 56, Appl
19	62	100.0	199	1	PCT-US01-09616-60	Sequence 60, Appl
20	62	100.0	199	22	US-09-815-847-56	Sequence 56, Appl
21	62	100.0	199	22	US-09-815-847-60	Sequence 60, Appl
22	62	100.0	208	1	PCT-US01-09616-102	Sequence 102, App
23	62	100.0	208	22	US-09-815-847-102	Sequence 102, App
24	62	100.0	209	1	PCT-US01-09616-58	Sequence 58, Appl
25	62	100.0	209	22	US-09-815-847-58	Sequence 58, Appl
26	62	100.0	210	1	PCT-US01-09616-40	Sequence 40, Appl
27	62	100.0	210	22	US-09-815-847-40	Sequence 4 Appl
28	62	100.0	211	1	PCT-US01-09616-39	Sequence 39, Appl
29	62	100.0	211	1	PCT-US01-09616-59	Sequence 59, Appl
30	62	100.0	211	22	US-09-815-847-39	Sequence 39, Appl
31	62	100.0	211	22	US-09-815-847-59	Sequence 59, Appl
32	62	100.0	212	1	PCT-US01-09616-37	Sequence 37, Appl
33	62	100.0	212	1	PCT-US01-09616-38	Sequence 38, Appl
34	62	100.0	212	22	US-09-815-847-37	Sequence 37, Appl
35	62	100.0	212	22	US-09-815-847-38	Sequence 38, Appl
36	62	100.0	213	1	PCT-US01-09616-41	Sequence 41, Appl
37	62	100.0	213	1	PCT-US01-09616-43	Sequence 43, Appl
38	62	100.0	213	1	PCT-US01-09616-44	Sequence 44, Appl
39	62	100.0	213	1	PCT-US01-09616-57	Sequence 57, Appl
40	62	100.0	213	22	US-09-815-847-41	Sequence 41, Appl
41	62	100.0	213	22	US-09-815-847-43	Sequence 43, Appl
42	62	100.0	213	22	US-09-815-847-44	Sequence 44, Appl
43	62	100.0	213	22	US-09-815-847-57	Sequence 57, Appl
44	62	100.0	214	1	PCT-US01-09616-42	Sequence 42, Appl
45	62	100.0	214	22	US-09-815-847-42	Sequence 42, Appl

ALIGNMENTS

RESULT 1
US-08-985-898-1
; Sequence 1, Application US/08985898
; GENERAL INFORMATION:
; APPLICANT: Rogers, Peters PGA
; TITLE OF INVENTION: Method of preparing a monoclonal
; TITLE OF INVENTION: antibody (as amended)
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Alfred Nobel Patent Dept.
; STREET: 1400 Rockville Drive, Suite 206
; CITY: Rockville
; STATE: Maryland
; COUNTRY: US
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/985,898
; FILING DATE: 5-DEC-1997
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Sullivan, Michael G.
; REGISTRATION NUMBER: 35,377
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 301-948-7400
; TELEFAX: 301-948-9751

```
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  13 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    FRAGMENT TYPE:  internal
US 08 985 898 1


  Query Match             100.0%;  Score 6.2;  DB 14;  Length 13;
  Best Local Similarity   100.0%;  Pred. No. 0.00018;
  Matches   13;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 RSFTLASSFTGVG 13
           ||||||||||||| |
Db       1 RSFTLASSFTGVG 13


RESULT   2
US 09 158 300 6
; Sequence 6, Application US/09158300
; GENERAL INFORMATION:
;    APPLICANT:  A.M.H. Boots
;    APPLICANT:  G.F.M. Verheijden
;    APPLICANT:  E.S. Bos
;    TITLE OF INVENTION:  Novel Peptides derived from Anti-Idiotypic T Cells
;    TITLE OF INVENTION:  in Immunotherapy of Autoimmune Diseases
;    NUMBER OF SEQUENCES:  10
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Akzo Nobel Patent Department
;      STREET:  1300 Piccard Drive, Suite 206
;      CITY:  Rockville
;      STATE:  Maryland
;      COUNTRY:  USA
;      ZIP:  20850
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/158,300
;      FILING DATE:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/644,493
;      FILING DATE:  18-APR 1996
;      APPLICATION NUMBER:  US 08/619,645
;      FILING DATE:  25-MAR-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/EP95/04201
;      FILING DATE:  25-OCT 1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  NL 9420 1287
;      FILING DATE:  27-OCT 1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  NL 9520 0860
;      FILING DATE:  07-APR 1995
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (301) 258 5300
;      TELEFAX:  (301) 977 3847
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  13 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US 09 158 300 6


  Query Match             100.0%;  Score 6.2;  DB 15;  Length 13;
  Best Local Similarity   100.0%;  Pred. No. 0.00018;
  Matches   13;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 RSFTLASSFTGVG 13
           ||||||||||||| |
Db       1 RSFTLASSFTGVG 13


RESULT   3
US 09 405 745 1
; Sequence 1, Application US/09405745
; GENERAL INFORMATION:
;    APPLICANT:  Stoonbakkers BGA
;    TITLE OF INVENTION:  Method of preparing a monoclonal
;    TITLE OF INVENTION:  antibody (as amended)
;    NUMBER OF SEQUENCES:  3
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Akzo Nobel Patent Dept.
;      STREET:  1300 Piccard Drive, Suite 206
;      CITY:  Rockville
;      STATE:  Maryland
;      COUNTRY:  US
;      ZIP:  20850
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC DOS/MS DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30 (EPO)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/405,745
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/ 08/585,898
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Sullivan, Michael G.
;      REGISTRATION NUMBER:  35,377
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  301-948-7400
;      TELEFAX:  301-948-9751
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  13 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    FRAGMENT TYPE:  internal
US 09 405 745 1


  Query Match             100.0%;  Score 6.2;  DB 16;  Length 13;
  Best Local Similarity   100.0%;  Pred. No. 0.00018;
  Matches   13;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 RSFTLASSFTGVG 13
           ||||||||||||| |
Db       1 RSFTLASSFTGVG 13


RESULT   4
US 09 744 170 3
; Sequence 3, Application US/09744170
; GENERAL INFORMATION:
;  APPLICANT: VERHEIJDEN, Gijsbertus F.M.
;  APPLICANT:  BOOTS, Anna M.H.
;  TITLE OF INVENTION: NOVEL PEPTIDES FOR USE IN IMMUNOTHERAPY I ACT IMMUNE
;  TITLE OF INVENTION:  DISEASES
;  FILE REFERENCE: VERHEIJDEN ET AL.
;  CURRENT APPLICATION NUMBER: US/09/744,170
;  CURRENT FILING DATE:  2001-04-05
;  PRIOR APPLICATION NUMBER: EP/EP99/05050
```

```
;  PRIOR FILING DATE: 1996-07-19
;  PRIOR APPLICATION NUMBER: EP 96,202,471.8
;  PRIOR FILING DATE: 1996-07-23
;  NUMBER OF SEQ ID NOS: 9
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO: 6
;   LENGTH: 13
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-744-282-6


  Query Match             100.0%;  Score 62;  DB 21;  Length 13;
  Best Local Similarity   100.0%;  Pred. No. 0.00018;
  Matches   13;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 RSFTLASSEIGVG 13
           |||||||||||||
Db       1 RSFTLASSEIGVG 13


RESULT   5
US-09-744-282-6
; Sequence 6, Application US/09744282
; GENERAL INFORMATION:
;  APPLICANT: MILLENBERG, Andreas M.H.
;  APPLICANT:  BOERS, Anne M.H.
;  TITLE OF INVENTION: USE OF BOVINE IN IMMUNE DISEASES
;  FILE REFERENCE: MILLENBERG
;  CURRENT APPLICATION NUMBER: US/09/744,282
;  CURRENT FILING DATE:  2001-04-09
;  PRIOR APPLICATION NUMBER: PCT/EP99/05331
;  PRIOR FILING DATE: 1999-07-19
;  PRIOR APPLICATION NUMBER: EP 98202471.8
;  PRIOR FILING DATE: 1998-07-23
;  NUMBER OF SEQ ID NOS: 9
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO: 6
;   LENGTH: 13
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-744-282-6


  Query Match             100.0%;  Score 62;  DB 21;  Length 13;
  Best Local Similarity   100.0%;  Pred. No. 0.00018;
  Matches   13;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 RSFTLASSEIGVG 13
           |||||||||||||
Db       1 RSFTLASSEIGVG 13


RESULT   6
US-09-985-065-1
; Sequence 1, Application US/09985065
;   GENERAL INFORMATION:
;        APPLICANT: Steenbakkers PGA
;        TITLE OF INVENTION: Method of preparing a monoclonal
;                            antibody (as amended)
;        NUMBER OF SEQUENCES: 4
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Akzo Nobel Patent Dept.
;             STREET: 1300 Piccard Drive, Suite 206
;             CITY: Rockville
;             STATE: Maryland
;             COUNTRY: US
;             ZIP: 20850
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/985,065
;             FILING DATE: 01-Nov-2001
;             CLASSIFICATION: Unknown
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 09/405,745
;             FILING DATE: 01-NOV-2001
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Sullivan, Michael G.
;             REGISTRATION NUMBER: 35,477
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 301-948-7400
;             TELEFAX: 301-948-9751
;   INFORMATION FOR SEQ ID NO: 1:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 13 amino acids
;             TYPE: amino acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        MOLECULE TYPE: protein
;        FRAGMENT TYPE: internal
;        SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-985-065-1


  Query Match             100.0%;  Score 62;  DB 23;  Length 13;
  Best Local Similarity   100.0%;  Pred. No. 0.00018;
  Matches   13;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 RSFTLASSEIGVG 13
           |||||||||||||
Db       1 RSFTLASSEIGVG 13


RESULT   7
US-08-985-898-2
; Sequence 2, Application US/08985898
;  GENERAL INFORMATION:
;    APPLICANT:  Steenbakkers PGA
;    TITLE OF INVENTION:  Method of preparing a monoclonal
;    TITLE OF INVENTION:  antibody (as amended)
;    NUMBER OF SEQUENCES:  4
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Akzo Nobel Patent Dept.
;      STREET:  1300 Piccard Drive, Suite 206
;      CITY:  Rockville
;      STATE:  Maryland
;      COUNTRY:  US
;      ZIP:  20850
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30 (EPO)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/985,898
;      FILING DATE:  5-DEC-1997
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Sullivan, Michael G.
;      REGISTERATION NUMBER:  35,477
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  301-948-7400
;      TELEFAX:  301-948-9751
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  15 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    FRAGMENT TYPE:  internal
US-08-985-898-2
```

```
  Query Match             100.0%;  Score 62;  DB 15;  Length 15;
  Best Local Similarity   100.0%;  Pred. No. 0.00021;
  Matches   13;  Conservative    0;  Mismatches    0;  Indels     0;  Gaps     0;

Qy       1 RSFTLASSETGVG 13
           |||||||||||||
Db       3 RSFTLASSETGVG 15


RESULT   8
US-09-405-745-2
; Sequence 2, Application US/09405745
;  GENERAL INFORMATION:
;     APPLICANT:  Stoonbakkers PGA
;     TITLE OF INVENTION:  Method of preparing a monoclonal
;     TITLE OF INVENTION:  antibody (as amended)
;     NUMBER OF SEQUENCES:  4
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Akzo Nobel Patent Dept.
;       STREET:  1300 Piccard Drive, Suite 206
;       CITY:  Rockville
;       STATE:  Maryland
;       COUNTRY:  US
;       ZIP:  20850
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.30 (EPO)
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/09/405,745
;       FILING DATE:
;       CLASSIFICATION:
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/985,898
;       FILING DATE:
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Sullivan, Michael G.
;       REGISTRATION NUMBER:  45,477
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  301-948-7160
;       TELEFAX:  301-948-9751
;   INFORMATION FOR SEQ ID NO:  2:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  15 amino acids
;       TYPE:  amino acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  protein
;     FRAGMENT TYPE:  internal
US-09-405-745-2


  Query Match             100.0%;  Score 62;  DB 18;  Length 15;
  Best Local Similarity   100.0%;  Pred. No. 0.00021;
  Matches   13;  Conservative    0;  Mismatches    0;  Indels     0;  Gaps     0;

Qy       1 RSFTLASSETGVG 13
           |||||||||||||
Db       3 RSFTLASSETGVG 15


RESULT   9
US-09-985-065-2
; Sequence 2, Application US/09985065
;  GENERAL INFORMATION:
;     APPLICANT:  Stoonbakkers PGA
;     TITLE OF INVENTION:  Method of preparing a monoclonal
;                          antibody (as amended)
;     NUMBER OF SEQUENCES:  4
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Akzo Nobel Patent Dept.
;       STREET:  1300 Piccard Drive, Suite 206
;       CITY:  Rockville
;       STATE:  Maryland
;       COUNTRY:  US
;       ZIP:  20850
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.30 (EPO)
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/09/985,065
;       FILING DATE:  01-NOV-2001
;       CLASSIFICATION:  424
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  09/405,745
;       FILING DATE:  01-NOV-2001
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Sullivan, Michael G.
;       REGISTRATION NUMBER:  45,477
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  301-948-7160
;       TELEFAX:  301-948-9751
;   INFORMATION FOR SEQ ID NO:  2:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  15 amino acids
;       TYPE:  amino acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  protein
;     FRAGMENT TYPE:  internal
;     SEQUENCE DESCRIPTION:  SEQ ID NO:  2:
US-09-985-065-2


  Query Match             100.0%;  Score 62;  DB 25;  Length 15;
  Best Local Similarity   100.0%;  Pred. No. 0.00021;
  Matches   13;  Conservative    0;  Mismatches    0;  Indels     0;  Gaps

Qy       1 RSFTLASSETGVG 13
           |||||||||||||
Db       3 RSFTLASSETGVG 15


RESULT   10
US-08-985-898-3
; Sequence 3, Application US/08985898
;  GENERAL INFORMATION:
;     APPLICANT:  Stoonbakkers PGA
;     TITLE OF INVENTION:  Method of preparing a monoclonal
;     TITLE OF INVENTION:  antibody (as amended)
;     NUMBER OF SEQUENCES:  4
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Akzo Nobel Patent Dept.
;       STREET:  1300 Piccard Drive, Suite 206
;       CITY:  Rockville
;       STATE:  Maryland
;       COUNTRY:  US
;       ZIP:  20850
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.30 (EPO)
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/985,898
;       FILING DATE:  5-DEC-1997
;       CLASSIFICATION:
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Sullivan, Michael G.
;       REGISTRATION NUMBER:  45,477
;     TELECOMMUNICATION INFORMATION:
```

```
;       TELEPHONE:  3   94- 7400
;       TELEFAX:  301 948 9751
;    INFORMATION FOR SEQ ID NO:  3:
;      SEQUENCE CHARACTERISTICS:
;         LENGTH:  16 amino acids
;         TYPE:  amino acid
;         STRANDEDNESS:  single
;         TOPOLOGY:  linear
;      MOLECULE TYPE:  protein
;      FRAGMENT TYPE:  internal
US-09-405-745-3


  Query Match             100.0%;  Score 62;  DB 18;  Length 16;
  Best Local Similarity   100.0%;  Prod. No. 0.00022;
  Matches   13;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 RSPTLASSETGVG 13
           |||||||||||||
Db       2 RSPTLASSETGVG 14


RESULT  11
US-09-405-745-4
; Sequence 4, Application US/09405745
;  GENERAL INFORMATION:
;     APPLICANT:  Steenbakkers PGA
;     TITLE OF INVENTION:  Method of preparing a monoclonal
;     TITLE OF INVENTION:  antibody (as amended)
;     NUMBER OF SEQUENCES:  4
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Akzo Nobel Patent Dept.
;       STREET:  1300 Piccard Drive, Suite 206
;       CITY:  Rockville
;       STATE:  Maryland
;       COUNTRY:  US
;       ZIP:  20850
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.30 (EPO)
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/09/405,745
;       FILING DATE:
;       CLASSIFICATION:
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US/09/405,745
;       FILING DATE:
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Sullivan, Michael G.
;       REGISTRATION NUMBER:  35,377
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  301-948-7400
;       TELEFAX:  301-948-9751
;   INFORMATION FOR SEQ ID NO:  3:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  16 amino acids
;       TYPE:  amino acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  protein
;     FRAGMENT TYPE:  internal
US-09-405-745-3


  Query Match             100.0%;  Score 62;  DB 18;  Length 16;
  Best Local Similarity   100.0%;  Prod. No. 0.00022;
  Matches   13;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 RSPTLASSETGVG 13
           |||||||||||||
Db       2 RSPTLASSETGVG 14


RESULT  12
US-09-985-065-3
; Sequence 3, Application US/09985065
;   GENERAL INFORMATION:
;         APPLICANT: Steenbakkers PGA
;         TITLE OF INVENTION: Method of preparing a monoclonal
;                             antibody (as amended)
;         NUMBER OF SEQUENCES: 4
;         CORRESPONDENCE ADDRESS:
;              ADDRESSEE: Akzo Nobel Patent Dept.
;              STREET: 1300 Piccard Drive, Suite 206
;              CITY: Rockville
;              STATE: Maryland
;              COUNTRY: US
;              ZIP: 20850
;         COMPUTER READABLE FORM:
;              MEDIUM TYPE: Floppy disk
;              COMPUTER: IBM PC compatible
;              OPERATING SYSTEM: PC-DOS/MS-DOS
;              SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
;         CURRENT APPLICATION DATA:
;              APPLICATION NUMBER: US/09/985,065
;              FILING DATE: 01-Nov-2001
;              CLASSIFICATION: <Unknown>
;         PRIOR APPLICATION DATA:
;              APPLICATION NUMBER: 09/405,745
;              FILING DATE: 01-NOV-2001
;         ATTORNEY/AGENT INFORMATION:
;              NAME: Sullivan, Michael G.
;              REGISTRATION NUMBER: 35,377
;         TELECOMMUNICATION INFORMATION:
;              TELEPHONE: 301-948-7400
;              TELEFAX: 301-948-9751
;    INFORMATION FOR SEQ ID NO: 3:
;         SEQUENCE CHARACTERISTICS:
;              LENGTH: 16 amino acids
;              TYPE: amino acid
;              STRANDEDNESS: single
;              TOPOLOGY: linear
;         MOLECULE TYPE: protein
;         FRAGMENT TYPE: internal
;         SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-09-985-065-3


  Query Match             100.0%;  Score 62;  DB 18;  Length 16;
  Best Local Similarity   100.0%;  Prod. No. 0.00022;
  Matches   13;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 RSPTLASSETGVG 13
           |||||||||||||
Db       2 RSPTLASSETGVG 14


RESULT  13
US-60-160-189-9314
; Sequence 9314, Application US/60160189
; GENERAL INFORMATION:
;  APPLICANT: BONAZZI, VIVIEN
;  TITLE OF INVENTION: ISOLATED HUMAN DRUG TARGET PROTEINS,
;  TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN DRUG TARGET PROTEINS
;  TITLE OF INVENTION: AND USES THEREOF
;  FILE REFERENCE: CL000112
;  CURRENT APPLICATION NUMBER: US/60/160,189
;  CURRENT FILING DATE: 1999-10-19
;  NUMBER OF SEQ ID NOS: 10162
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 9314
;   LENGTH: 34
;   TYPE: PRT
;   ORGANISM: HUMAN
```

US-60-160-203-5456

  Query Match 100.0%; Score 62; DB 26; Length 44;
  Best Local Similarity 100.0%; Pred. No. 0.00054;
  Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RSFILASSEIGVG 13
|||||||||||||
Db 13 RSFILASSEIGVG 25

RESULT 14
US-60-160-203-5456
; Sequence 5456, Application US/60160203
; GENERAL INFORMATION:
; APPLICANT: BONAZZI, VIVIEN
; TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: PI1000116
; CURRENT APPLICATION NUMBER: US/60/160,203
; CURRENT FILING DATE: 1999-10-19
; NUMBER OF SEQ ID NOS: 6474
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 5456
; LENGTH: 44
; TYPE: PRT
; ORGANISM: HUMAN
US-60-160-203-5456

  Query Match 100.0%; Score 62; DB 26; Length 44;
  Best Local Similarity 100.0%; Pred. No. 0.00073;
  Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RSFILASSEIGVG 13
||-|-|||-|||
Db 23 RSFTLASSEIGVG 35

RESULT 15
US-60-169-840-6915
; Sequence 6915, Application US/60169840
; GENERAL INFORMATION:
; APPLICANT: Bonazzi, Vivien
; TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: PI1000164
; CURRENT APPLICATION NUMBER: US/60/169,840
; CURRENT FILING DATE: 1999-12-09
; NUMBER OF SEQ ID NOS: 9028
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 6915
; LENGTH: 44
; TYPE: PRT
; ORGANISM: Human
US-60-169-840-6915

  Query Match 100.0%; Score 62; DB 26; Length 44;
  Best Local Similarity 100.0%; Pred. No. 0.00073;
  Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RSFILASSEIVG 13
||| |||||-|||
Db 23 RSFILASSEIVG 35

Search completed: August 26, 2002, 06:22:19
Job time: 899 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

M protein: protein search, using sw model

Run on: August 26, 2002, 05:08:54 ; Search time 8.5 Seconds
(without alignments)
10.743 Million cell updates/sec

Title: US-09-744-282-1
Perfect score: 44
Sequence: 1 FGRSFTLAS 9

Scoring table: BLOSUM62
Gap 10.0 , Gapext 0.5

Searched: 283138 seqs, 99999934 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 71:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	44	100.0	388	2	S51327	heparin-binding gl
2	44	100.0	388	2	A49562	cartilage glycopro
3	41	93.2	405	2	S61561	breast carcinoma
4	38	86.3	525	2	I44445	chitinase (EC 3.2.
5	36	81.8	203	2	G25511	Cp protein - trut
6	36	81.8	1074	2	F72217	conserved hypothet
7	35	79.5	554	2	A56596	chitinase (EC 3.2.
8	35	79.5	765	2	E84553	hypothetical prote
9	34	77.3	314	2	F89929	response regulator
10	33	75.0	267	2	I36271	conserved hypothet
11	33	75.0	684	2	F69408	conserved hypothet
12	33	75.0	1645	2	I14075	chitinase (EC 3.2.
13	32	72.7	114	2	G71048	hypothetical prote
14	32	72.7	307	2	E98174	oligopeptide ABC t
15	32	72.7	307	2	AB3112	hypothetical prote
16	32	72.7	333	2	S56753	protochlorophyllid
17	32	72.7	333	2	S37823	protochlorophyllid
18	32	72.7	333	2	S30915	protochlorophyllid
19	32	72.7	334	2	I50882	chlorophyllide red
20	32	72.7	405	2	I39922	hypothetical prote
21	32	72.7	415	2	S54140	DNA binding protei
22	32	72.7	415	2	S57690	DNA binding protei
23	32	72.7	489	2	A59192	chitinase (EC 3.2.
24	32	72.7	592	2	I41078	serine/threonine p
25	32	72.7	741	2	I44237	hypothetical prote
26	32	72.7	1033	2	I37715	actin-interacting
27	31	70.5	246	2	I20341	hypothetical prote
28	31	70.5	315	2	C70899	probable didipeptid
29	31	70.5	425	2	I39636	ring finger protei
30	31	70.5	565	2	I38941	DNA directed RNA p
31	31	70.5	573	2	JQ0135	hypothetical 57.8K
32	31	70.5	716	2	AC2449	ABC transporter AT
33	31	70.5	1733	2	S27939	tensin - chicken
34	31	70.5	1744	2	A54970	tensin, cardiac mu
35	31	70.5	1792	2	A57075	tensin - chicken (
36	30	68.2	114	2	F82148	hypothetical prot
37	30	68.2	179	2	AG1994	hypothetical prote
38	30	68.2	196	2	H72551	hypothetical prote
39	30	68.2	199	1	S50398	ribosomal protein
40	30	68.2	199	2	S67618	ribosomal protein
41	30	68.2	202	2	I52146	ribosomal protein
42	30	68.2	292	2	S77168	hypothetical prote
43	30	68.2	338	2	AE3079	hypothetical prote
44	30	68.2	376	2	G95949	probable oligopept
45	30	68.2	381	2	AB0734	probable bacteriop

ALIGNMENTS

RESULT 1
S51327
heparin-binding glycoprotein 38K - pig
C;Species: Sus scrofa domestica (domestic pig)
C;Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 22-Jun-1999
C;Accession: S51327
R;Shackelton, L.M.; Mann, D.M.; Millis, A.J.T.
submitted to the EMBL Data Library, January 1995
A;Description: Identification of a 38kDa heparin binding glycoprotein (gp38k) in adult
A;Reference number: S51327
A;Accession: S51327
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-383 <SHA>
A;Cross-references: EMBL:Z47803; NID:g634097; PIDN:CAA87764.1; PID:g634098
C;Superfamily: Streptomyces chitinase chi40

Query Match 100.0%; Score 44; DB 2; Length 383;
Best Local Similarity 100.0%; Pred. No. 0.14;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FGRSFTLAS 9
|||||||||
Db 261 FGRSFTLAS 269

RESULT 2
A49562
cartilage glycoprotein gp39 precursor - human
N;Alternate names: 39K synovial protein
C;Species: Homo sapiens (man)
C;Date: 23-Mar-1995 #sequence_revision 23-Mar-1995 #text_change 22-Jun-1999
C;Accession: A49562; S10677; A33162
R;Hakala, B.E.; White, C.; Recklies, A.D.
J. Biol. Chem. 268, 25803-25810, 1993
A;Title: Human cartilage gp-39, a major secretory product of articular chondrocytes a
A;Reference number: A49562; MUID:94064658
A;Accession: A49562
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-383 <HAK>
A;Cross-references: GB:M80927; NID:g348911; PIDN:AAA16074.1; PID:g348912
R;Nyirkos, P.; Golds, E.E.
Biochem. J. 269, 265-268, 1990
A;Title: Human synovial cells secrete a 39 kDa protein similar to a bovine mammary pr
A;Reference number: S10677; MUID:90328983
A;Accession: S10677
A;Molecule type: protein
A;Residues: 22-40,'X',42-45 <NY2>
C;Superfamily: Streptomyces chitinase chi40
C;Keywords: cartilage; extracellular protein; glycoprotein

F;1 21/Domain: signal sequence #status predicted -SIG-
F;22-383/Product: carbohydrate glycoprotein gp39 #status predicted: MAT-

Query Match 100.0%; Score 44; DB 2; Length 98;
Best Local Similarity 100.0%; Pred. No. 0.14;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PGRSFILAS 9
||||||||||
Db 261 PGRSFILAS 269

RESULT 3
S61551
breast regressina protein brp39 precursor - mouse
C;Species: Mus musculus (house mouse)
C;Date: 19-Mar-1997 #sequence_revision 25-Apr-1997 #text_change 23-Jun-2000
C;Accession: S61551; S61550; 148271
R;Morrison, B.W.; Leder, P.
oncogene 9, 3417-3426, 1994
A;Title: neu and ras initiate murine mammary tumors that share genetic markers generally
A;Reference number: 148271; MUID:95006797
A;Accession: S61551
A;Status: nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1-405 <MOK1>
A;Cross-references: EMBL:X93035; NID:g1085065; PIDN:CAA63603.1; PID:g1085066
R;Morrison, B.W.
submitted to the EMBL Data Library, November 1996
A;Reference number: S61550
A;Accession: S61550
A;Molecule type: mRNA
A;Residues: 1-347,<347-433,<435-453,<455-*MVWAIFFITLQLVVSAAHSLSCAGA* MOB2-
A;Cross-references: EMBL:X93035; NID:g1085065; PIDN:CAA63603.1; PID:g1085066
A;Note: the differences at the carboxyl end are due to a frameshift error
C;Genetics:
A;Gene: brp39
C;Superfamily: Streptomyces chitinase chi40
F;1-21/Domain: signal sequence #status predicted -SIG-
F;22-405/Product: breast regressing protein brp39 #status predicted -MAT-

Query Match 93.2%; Score 41; DB 2; Length 105;
Best Local Similarity 88.9%; Pred. No. 0.66;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PGRSFILAS 9
|||:||||||
Db 262 PGKSFILAS 270

RESULT 4
I44445
chitinase (EC 3.2.1.14) [imported] - African malaria mosquito
C;Species: Anopheles gambiae (African malaria mosquito)
C;Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 21-Jan-2000
C;Accession: I44445
R;Shen, Z.; Jacobs-Lorena, M.
submitted to the EMBL Data Library, June 1997
A;Reference number: Z22721
A;Accession: I44445
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-525 <SHE>
A;Cross-references: EMBL:AF008575; PIDN:AAB82764.1
A;Experimental source: adult; gut
C;Genetics:
A;Gene: chi-1
C;Keywords: glycosidase; hydrolase

Query Match 86.4%; Score 38; DB 2; Length 327;
Best Local Similarity 77.8%; Pred. No. 4.9;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PGRSFILAS 9
:.|:|||||
Db 279 YGRNFILAS 287

RESULT 5
C29511
C;Protein - fruit fly (Drosophila melanogaster)
C;Species: Drosophila melanogaster
C;Date: 30-Jun-1988 #sequence_revision 30-Jun-1988 #text_change 27-[illegible]-2000
C;Accession: C29511
R;Ezeloth Jr., D.E.; Marsh, J.L.
Nucleic Acids Res. 14, 6195-6196, 1986
A;Title: Sequence and expression of the 5' gene, a member of the [illegible] gene family
A;Reference number: A29634; MUID:86312087
A;Accession: C29511
A;Molecule type: DNA
A;Residues: 1-204 <EVE>
A;Cross-references: GB:X04227; NID:g769; PIDN:CAA28[illegible]; PID:g[illegible]
C;Genetics:
A;Gene: FlyBase:[illegible]
A;Cross-references: FlyBase:FBgn[illegible]
A;Introns: 67/3

Query Match 81.8%; Score 36; DB 2; Length [illegible];
Best Local Similarity 87.5%; Pred. No. [illegible];
Matches 7; Conservative 1; Mismatches 1; Indels [illegible]; Gaps [illegible];

Qy 1 PGRSFILA 8
||| ||||
Db 98 PGKEFILA 105

RESULT 6
F72217
conserved hypothetical protein - Thermotoga maritima (strain MSB8)
C;Species: Thermotoga maritima
C;Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 21-[illegible]-2000
C;Accession: F72217
R;Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, [illegible]
Garrett, M.M.; Stewart, A.M.; Patton, M.D.; Pratt, M.P.; Ferguson, [illegible]
C.M.
Nature 399, 323-329, 1999
A;Title: Evidence for lateral gene transfer between Archaea and bacteria from genome
A;Reference number: A72200; MUID:99287316
A;Accession: F72217
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1074 <NEL>
A;Cross-references: GB:AE000[illegible]; GB:AF0[illegible]; NID:g4[illegible]
A;Experimental source: strain MSB8
C;Genetics:
A;Gene: TM1736

Query Match 81.8%; Score 36; DB 2; Length [illegible];
Best Local Similarity 77.8%; Pred. No. [illegible];
Matches 7; Conservative 1; Mismatches 1; Indels [illegible]; Gaps [illegible];

Qy 1 PGRSFILAS 9
|||:||| |
Db 1038 PGRNFILES 1046

RESULT 7
A66596
chitinase (EC 3.2.1.14) - tobacco hornworm
C;Species: Manduca sexta (tobacco hornworm)
C;Date: 11-Aug-1995 #sequence_revision 21-Aug-1998 #text_change [illegible]-Jan-2000

C;Accession: A56590
R;Kramer, K.J.; Corpuz, L.; Choi, H.K.; Muthukrishnan, S.
Insect Biochem. Mol. Biol. 23, 691-701, 1993
A;Title: Sequence of a cDNA and expression of the gene encoding epidermal and gut chitin
A;Reference number: A56590; MUID:93357793
A;Accession: A56590
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-554 <KRA>
A;Cross-references: GB:L22270; GB:S64757; NID:g400148; PID:g46-49
A;Experimental source: silkworm larvae
A;Note: sequence extracted from NCBI backbone (NCBIN:106417, NCBIP:106418)
C;Keywords: glycosidase; hydrolase; polysaccharide degradation

Query Match 79.5%; Score 45; DB 2; Length 554;
Best Local Similarity 66.7%; Pred. No. 17;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FGRSFTLAS 9
:|:|||||:
Db 270 YGRSFTLSA 278

RESULT 8
E84553
hypothetical protein At2g17550 [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse ear cress)
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
C;Accession: E84553
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; M.; Kee, H.; Moffat, K.S.; Unseld, M.J.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L. ouss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J Nature 402, 761-768, 1999
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A;Reference number: A84120; MUID:20034487
A;Accession: E84553
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-765 <LIN>
A;Cross-references: GB:AC002560; NID:g3021276; PID:AAC02613.1; GSPDB:GN00126
C;Genetics:
A;Gene: At2g17550
A;Map position: 2

Query Match 76.5%; Score 43; DB 2; Length 765;
Best Local Similarity 77.8%; Pred. No. 24;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 FGRSFTLAS 9
:|| ||| |
Db 211 FGKEFTLKS 219

RESULT 9
B69629
response regulatory protein (Sensory transduction system) BH1284 [imported] - Bacillus h
C;Species: Bacillus halodurans
C;Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001
C;Accession: B69629
R;Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fuji, F.; Hir Nucleic Acids Res. 28, 4317-4331, 2000
A;Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A;Reference number: A69629; MUID:20512682; PMID:11058126
A;Accession: B69629
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-814 <TAK>
A;Cross-references: PIR:A69629; GB:BA000004; NID:g10174511; PID:BAA83573.1; GSPDB:GN
A;Experimental source: strain C-125
C;Genetics:
A;Gene: BH1284

C;Superfamily: Synechocystis sensory transduction system regulatory protein slr1759 ;

Query Match 77.3%; Score 34; DB 2; Length 414;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 GRSFTLA 8
|||:|||
Db 181 GRSFTLA 187

RESULT 10
F82701
conserved hypothetical protein XF1272 [imported] - Xylella fastidiosa (strain 9a5c)
C;Species: Xylella fastidiosa
C;Date: 18-Aug-2000 #sequence_revision 28-Aug-2000 #text_change 28-Aug-2000
C;Accession: F82701
R;Anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Seq Nature 406, 151-157, 2000
A;Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A;Reference number: A82515; MUID:20465717
A;Note: for a complete list of authors see reference number A59428 below
A;Accession: F82701
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-267 <SIM>
A;Cross-references: GB:AE003961; GB:AE003849; NID:g9106254; PID:AAF84281.1; GSPDB:GN
A;Experimental source: strain 9a5c
R;Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R. Briones, M.R.S.; Bucno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer as-Neto, E.; Dorena, C.; El-Dorry, H.; Facincani, A.P.; Ferreira, A.J.S. submitted to GenBank, June 2000
A;Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Fr J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; La chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins A;Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C. , E.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Pereira, Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, E.V.; Sawa A;Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Sil M.; Tsukada, M.H.; Valdia, B.; Van Sluys, M.A.; Vencio, R.Z.; de Almeida, S.; Vettore, A.L. A;Reference number: A59428
A;Contents: annotation
C;Genetics:
A;Gene: XF1272

Query Match 75.0%; Score 33; DB 2; Length 247;
Best Local Similarity 85.7%; Pred. No. 24;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FGRSFTL 7
:||||||
Db 126 YGRSFTL 132

RESULT 11
B69308
conserved hypothetical protein AF0466 - Archaeoglobus fulgidus
C;Species: Archaeoglobus fulgidus
C;Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 22-Oct-1999
C;Accession: B69308
R;Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dod , Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.C.; Weidman, J.F.; McDonald, L. Nature 390, 364-370, 1997
A;Authors: Otterback, L.; Cotton, M.D.; Springs, L.; Artiach, L.; Kaine, B.P.; Sypes, Smith, H.O.; Woese, C.R.; Venter, J.C.
A;Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing arch
A;Reference number: A61250; MUID:98049344
A;Accession: B69308
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA

A;Residues: 1-684 <KLE>
A;Cross-references: GB:AE001072; GB:AE000782; NID:g2689095; PID:AAB90771.1; PID:g2650

Query Match 75.0%; Score 33; DB 2; Length 684;
Best Local Similarity 77.8%; Pred. No. 58;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 FGRSFILAS 9
|||||||||
Db 496 FGLGFILAS 504

RESULT 12
I14075
chitinase (EC 3.2.1.14) - yellow fever mosquito
C;Species: Aedes aegypti (yellow fever mosquito)
C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 21-Sep-1999
C;Accession: I14075
R;de la Vega, H.; Specht, C.A.; Liu, Y.; Robbins, P.W.
Insect Mol. Biol. 7, 233-239, 1997
A;Title: Chitinases are a multi-gene family in Aedes, Anopheles and Drosophila.
A;Reference number: Z17872
A;Accession: I14075
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-535 <DEL>
A;Cross-references: EMBL:AF020192; NID:g2564729; PID:AAC83195.1
C;Genetics:
A;Gene: CHI
A;Function: 35,77-524,77; 578,81-603,91; 655,92
C;Keywords: glycosidase; hydrolase; polysaccharide degradation

Query Match 75.0%; Score 33; DB 2; Length 684;
Best Local Similarity 75.0%; Pred. No. 1.4e-02;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FGRSFILA 8
|:|||||.
Db 488 YGQSFILA 495

RESULT 13
D71048
hypothetical protein PH1676 - Pyrococcus horikoshii
C;Species: Pyrococcus horikoshii
C;Date: 14-Aug-1998 #sequence_revision 14-Aug-1998 #text_change 21-Jul-2000
C;Accession: D71048
R;Kawarabayasi, Y.; Sawada, M.; Horikawa, H.; Haikawa, Y.; Hino, Y.; Yamamoto, S.; Sek
M.; Ohtsuka, W.; Roppongi, I.; Uenaka, T.; Matoba, Y.; Yoshitake, N.; Kos
DNA Res. 5, 55-76, 1998
A;Title: Complete sequence and gene organization of the genome of a hyper-thermophilic
A;Reference number: A71000; MUID:98344147
A;Accession: D71048
A;Status: preliminary; nucleic acid sequence not shown; translation not shown.
A;Molecule type: DNA
A;Residues: 1-114 <KAW>
A;Cross-references: GB:AP000006; NID:g4235174; PID:BAA30732.1; PID:g4581
A;Experimental source: strain OT3
A;Note: this accession replaces an interim accession for a sequence replaced by Genbank
C;Genetics:
A;Gene: PH1676

Query Match 74.7%; Score 32; DB 2; Length 114;
Best Local Similarity 77.8%; Pred. No. 16;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FGRSFILAS 9
|:|||||||
Db 1 FGKSFSLAS 9

RESULT 14
E98174
oligopeptide ABC transporter (permease) (AP001508) [imported] - Agrobacterium tumefaci
C;Species: Agrobacterium tumefaciens
C;Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 11-Jan-2002
C;Accession: E98174
R;Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldm
A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Mark
Science 294, 2323-2328, 2001
A;Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium
A;Reference number: A97359; PMID:11743194
A;Accession: E98174
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-307 <KUK>
A;Cross-references: GB:AE008787; PIDN:AAK89070.1; PID:g15160655; GB:AP002508; NID:g
C;Genetics:
A;Gene: AGR_L_689
A;Map position: linear chromosome

Query Match 74.7%; Score 32; DB 2; Length 307;
Best Local Similarity 100.0%; Pred. No. 47;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FGRSFI 6
||||||
Db 78 FGRSFI 83

RESULT 15
AB3112
hypothetical protein Atu4524 [imported] - Agrobacterium tumefaciens (strain C58, Cer
C;Species: Agrobacterium tumefaciens
C;Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 11-Jan-2002
C;Accession: AB3112
R;Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Weed, T.; Thomas, V.; W
erade, G.; Gillet, W.; Grant, C.; Guenthner, L.; Kutyavin, I.; Lee, S.; Lu, M.; Mo
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A;Authors: Yoo, H.; Jan, Y.; Biddle, P.; Jones, M.; Krospan, W.; Lee, M.; Lubin, Ka
ster, E.W.
A;Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A;Reference number: AB2577; PMID:11743193
A;Accession: AB3112
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-307 <KUK>
A;Cross-references: GB:AE008688; PIDN:AAL42496.1; PID:g15160655; GB:AP002508; NID:g
A;Experimental source: strain C58 (Dupont)
C;Genetics:
A;Gene: Atu4524
A;Map position: linear chromosome

Query Match 74.7%; Score 32; DB 2; Length 307;
Best Local Similarity 100.0%; Pred. No. 47;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FGRSFI 6
||||||
Db 78 FGRSFI 83

Search completed: August 26, 2002, 06:10:44
Job time: 105 sec

Geneticx version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

SM protein - protein search, using sw model

Run on: August 26, 2002, 06:14:34 ; Search time 45.99 Seconds
(without alignments)
4,583 Million cell updates/sec

Title: US-09-744-282-1
Perfect score: 44
Sequence: 1 EQESFLASQ 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105,224 seqs, 38729550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_43:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	44	100.0	383	1	CH3L1_HUMAN	P36222 homo sapien
2	41	93.2	381	1	CH3L1_MOUSE	Q61362 mus musculu
3	36	81.8	403	1	LAMC_DROME	P21156 drosophila
4	36	81.8	760	1	ATL2_HUMAN	Q15782 homo sapien
5	35	79.5	554	1	BH11_MANSE	P81534 bombina bom
6	32	72.7	433	1	BCHX_RHOCA	P26504 rhodobacter
7	32	72.7	433	1	BCHX_RHOSH	Q02161 rhodobacter
8	32	72.7	405	1	ZR1_CAEEL	Q27571 caenorhabdi
9	32	72.7	415	1	ANL2_MICSA	P80021 microptenus
10	32	72.7	1585	1	EA11_SCHPO	Q10345 schizosacch
11	31	70.5	525	1	AN36_HUMAN	P17029 homo sapien
12	31	70.5	1744	1	TENS_CHICK	Q04209 gallus gall
13	30	68.2	199	1	SIR4_YEAST	Q12509 saccharomyc
14	30	68.2	299	1	BIB6_YEAST	P40212 saccharomyc
15	30	68.2	402	1	BIL5_CANAL	Q60083 candida alb
16	30	68.2	449	1	AN36_HUMAN	P17029 homo sapien
17	30	68.2	669	1	YVE2_SYNY3	Q59377 synechocyst
18	30	68.2	718	1	YDC0_MYCTN	P73432 mycoplasma
19	29	65.9	174	1	SSB_KLEPN	P23202 klebsiella
20	29	65.9	518	1	PE1_EBV	P03183 epstein bar
21	29	65.9	576	1	BCA2_YEAST	P47176 saccharomyc
22	29	65.9	577	1	YSY_SULSO	P80148 sulfolobus
23	29	65.9	593	1	PE1A_EBVCH	P29155 cercopithec
24	29	65.9	606	1	ALN_MOUSE	Q06287 mus musculu
25	29	65.9	601	1	SEN3_MOUSE	P30796 mus musculu
26	29	65.9	602	1	SEN3_MOUSE	P30797 mus musculu
27	29	65.9	604	1	NSP3_EBVCH	P21704 cercopithec
28	29	65.9	678	1	YAS4_RHISN	P55441 rhizobium s
29	29	65.9	508	1	YBY1_YEAST	P38181 saccharomyc
30	29	65.9	559	1	SYL_PASMU	P57847 pasteurella
31	29	65.9	663	1	UVRB_AQUAE	O67706 aquifex ae
32	29	65.9	720	1	PDE8_HUMAN	Q92004 homo sapien
33	29	65.9	740	1	CYG4_RAT	Q62714 rattus norv
34	29	65.9	742	1	CYG4_HUMAN	P54 [illegible] sapien
35	29	65.9	772	1	CPTM_HUMAN	Q96 [illegible] sapien
36	29	65.9	772	1	CPTM_RAT	Q63784 rattus norv
37	29	65.9	794	1	YG2A_YEAST	P53343 saccharomyc
38	29	65.9	839	1	ZP95_HUMAN	Q99218 homo sapien
39	29	65.9	872	1	IP3L_RAT	P42835 rattus norv
40	29	65.9	1206	1	RPOC_BACHD	Q92 [illegible] bacillus ha
41	29	65.9	1335	1	XDH_MOUSE	Q00519 mus musculu
42	28	63.6	146	1	HDE1_CAPHI	P [illegible] capra hir
43	28	63.6	146	1	HDE_TARBA	[illegible] tarsius ban
44	28	63.6	147	1	HDE_TARSY	P [illegible] tarsius syr
45	28	63.6	159	1	YMH7_YEAST	Q [illegible] saccharomyc

ALIGNMENTS

RESULT 1
CH3L1_HUMAN
ID CH3L1_HUMAN STANDARD; PRT: 383 AA.
AC P36222; P30923;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Chitinase-3 like protein 1 precursor (Cartilage glycoprotein-39) (GP-
DE 39) (39 kDa synovial protein) (YKL-40).
GN CHI3L1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC TISSUE=Cartilage;
RX MEDLINE=94064658; PubMed=8245017;
RA Hakala B.E., White C., Recklies A.D.;
RT "Human cartilage gp-39, a major secretory product of articular
RT chondrocytes and synovial cells, is a mammalian member of a chitinase
RT protein family.";
RL J. Biol. Chem. 268:25803-25810(1993).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Blood;
RX MEDLINE=97396591; PubMed=9244440;
RA Rehli M., Krause S.W., Andreesen R.;
RT "Molecular characterization of the gene for human cartilage gp-39
RT (CHI3L1), a member of the chitinase protein family and marker for
RT late stages of macrophage differentiation.";
RL Genomics 43:221-225(1997).
RN [3]
RP SEQUENCE OF 22-45.
RX MEDLINE=90228987; PubMed=2475755;
RA Nyirkos P., Golds E.E.;
RT "Human synovial cells secrete a 39 kDa protein similar to a bovine
RT mammary protein expressed during the non-lactating period.";
RL Biochem. J. 269:265-268(1990).
CC ! FUNCTION: MAY PLAY AN IMPORTANT ROLE IN THE CAPACITY OF CELLS I
CC RESPOND TO AND COPE WITH CHANGES IN THEIR ENVIRONMENT.
CC ! SUBCELLULAR LOCATION: Extracellular.
CC ! TISSUE SPECIFICITY: PRESENT IN ARTICULAR CHONDROCYTES, SYNOVIAL
CC CELLS AS WELL AS IN LIVER, UNDIFFERENTIATED IN NORMAL TISSUES, LUNG,
CC PANCREAS, MONONUCLEAR CELLS, OR FIBROBLASTS.
CC ! PTM: GLYCOSYLATED.
CC ! SIMILARITY: BELONGS TO FAMILY 18 OF GLYCOSYL HYDROLASES.
CC ----------- ---- --------- -- -------- ----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC ----------- ---- --------- -- -------- ----

```
DR   EMBL; M80947; AAA16074.1; -.
DR   EMBL; Y08374; CAA69661.1; -.
DR   EMBL; Y08375; CAA69661.1; JOINED.
DR   EMBL; Y08376; CAA69661.1; JOINED.
DR   EMBL; Y08377; CAA69661.1; JOINED.
DR   EMBL; Y08378; CAA69661.1; JOINED.
DR   PIR; S10677; S10677.
DR   PIR; A33162; A33162.
DR   MIM; 601525; -.
DR   InterPro; IPR001579; Chitinase_2.
DR   InterPro; IPR001223; Glyco_hydro_18.
DR   Pfam; PF00704; Glyco_hydro_18; 1.
DR   PROSITE; PS01095; CHITINASE_18; FALSE_NEG.
KW   Glycoprotein; Signal.
FT   SIGNAL          1     21
FT   CHAIN          22    383       CHITINASE 4 LIKE PROTEIN 1.
FT   CARBOHYD       60     60       N-LINKED (GLCNAC...) (POTENTIAL).
SQ   SEQUENCE   383 AA;  42613 MW;  76AD8298EEEC2D1 CRC64;


  Query Match             100.0%;  Score 44;  DB 1;  Length 383;
  Best Local Similarity   100.0%;  Pred. No. 0.05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 FGRSFTLAS 9
           |||||||||
Db     262 FGRSFTLAS 270


RESULT   2
C4L1_MOUSE
ID   C4L1_MOUSE      STANDARD;      PRT:   381 AA.
AC   Q61362;
DT   16-OCT-1995 (Rel. 32, Created)
DT   16-OCT-1995 (Rel. 32, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Chitinase 4 like protein 1 precursor (Cartilage glycoprotein-39)
DE   (GP-39) (BRP39 protein).
GN   CHI3L1 OR BRP39.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN FVB/N; TISSUE=Breast;
RX   MEDLINE=95060767; PubMed=7920700;
RA   Morrison B.W., Leder P.;
RT   "hcm and fgf-8-initiated murine mammary tumors share molecular markers
RT   generally absent in c-myc and int-2 initiated tumors.";
RL   Oncogene 9:3417-3426(1994).
CC   -!- FUNCTION: MAY PLAY AN IMPORTANT ROLE IN THE ACTIVITY OF CELLS IN
CC       RESPONSE TO AND/OR WITH CHANGES IN THEIR ENVIRONMENT (BY
CC       SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: Extracellular (By similarity).
CC   -!- SIMILARITY: BELONGS TO FAMILY 18 OF GLYCOSYL HYDROLASES.
CC   -----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----------------------------------------------------------------------
DR   EMBL; X76035; CAA53603.1; -.
DR   MGD; MGI:1347866; Chi3l1.
DR   InterPro; IPR001579; Chitinase_2.
DR   InterPro; IPR001223; Glyco_hydro_18.
DR   Pfam; PF00704; Glyco_hydro_18; 1.
DR   PROSITE; PS01095; CHITINASE_18; FALSE_NEG.
KW   Glycoprotein; Signal.
FT   SIGNAL          1     21       POTENTIAL.
FT   CHAIN          22    381       CHITINASE 4 LIKE PROTEIN 1.
FT   CARBOHYD       60     60       N-LINKED (GLCNAC...) (POTENTIAL).
SQ   SEQUENCE   381 AA;  43031 MW;  EE65B1E814E0450 CRC64;


  Query Match             93.2%;  Score 41;  DB 1;  Length ???;
  Best Local Similarity   88.9%;  Pred. No. 0.22;
  Matches    8;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 FGRSFTLAS 9
           ||.||||||
Db     262 FGKSFTLAS 270


RESULT   3
L2CC_DROME
ID   L2CC_DROME      STANDARD;      PRT:   203 AA.
AC   P24156;
DT   01-MAR-1992 (Rel. 21, Created)
DT   01-MAR-1992 (Rel. 21, Last sequence update)
DT   01-OCT-1996 (Rel. 34, Last annotation update)
DE   L(2)37Cc protein.
GN   L(2)37Cc.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=86312887; PubMed=3092183;
RA   Eveleth D.D. Jr., Marsh J.L.;
RT   "Sequence and expression of the Dc gene, a member of the dopa
RT   decarboxylase gene cluster of Drosophila: possible implications for
RT   regulation.";
RL   Nucleic Acids Res. 14:6169-6183(1986).
CC   -!- FUNCTION: REQUIRED FOR LARVAL METABOLISM OR FOR THE EMERGENCE
CC       OF THE LARVA INTO A PUPA.
CC   -!- SIMILARITY: BELONGS TO THE PROHIBITIN FAMILY.
CC   -----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----------------------------------------------------------------------
DR   EMBL; X04228; CAA27816.1; -.
DR   EMBL; X04227; CAA27807.1; -.
DR   PIR; C25511; C25511.
DR   FlyBase; FBgn0002131; L(2)37Cc.
DR   InterPro; IPR001132; Band_7.
DR   Pfam; PF01145; Band_7; 1.
DR   SMART; SM00244; PAB; 1.
KW   Developmental protein.
SQ   SEQUENCE   203 AA;  22817 MW;  B15D085CC0BC2A11 CRC64;


  Query Match             91.8%;  Score 40;  DB 1;  Length ???;
  Best Local Similarity   87.5%;  Pred. No. 1.5;
  Matches    7;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 FGRSFTLAS 9
           |||.||||.
Db     98 FGKEFILA 105


RESULT   4
C4L2_HUMAN
ID   C4L2_HUMAN      STANDARD;      PRT:   390 AA.
AC   Q15782; Q15783; Q 5749;
DT   16-OCT-2001 (Rel. 40, Created)
```

```
ID   [illegible] 2001 (Rel. 41, Last sequence update)
ID   [illegible] 2001 (Rel. 41, Last annotation update)
DE   Chitinase 3 like protein 2 precursor (YKL-39) (Cartilage glycoprotein
DE   39).
GN   CHI3L2.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A., [illegible] AND [illegible] MS (LIVER).
RA   Grossman A., McCormick T., Baker E., Waterhouse P., Sutherland G.R.,
RA   Mak T.W.;
RT   "Cloning of a novel lymphoid restricted human chitinase and
RT   localization to 1p13.3.";
RL   Submitted (MAY-1996) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A. (CHONDROSARCOMA).
RC   TISSUE=Articular cartilage;
RX   MEDLINE=96325097; PubMed=8702654;
RA   Hu B., Trask K., Pineda W.F., Price P.A.;
RT   "Isolation and sequence of a novel human chondrocyte protein related
RT   to mammalian members of the chitinase protein family.";
RL   J. Biol. Chem. 271:19415-19420(1996).
CC   -!- ALTERNATIVE PRODUCTS; 2 ISOFORMS; A SHORT FORM (CHI3L2-S) AND A
CC       LONG FORM; MAY BE PRODUCED BY ALTERNATIVE SPLICING.
CC   -!- TISSUE SPECIFICITY: HIGHEST EXPRESSION IN CHONDROSARCOMA, FOLLOWED
CC       BY SYNOVIOCYTES, LUNG AND BRAIN. NOT DETECTED IN BRAIN, SKELETAL
CC       PANCREAS, AND LIVER.
CC   -!- SIMILARITY: BELONGS TO FAMILY 18 OF GLYCOSYL HYDROLASES.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; U58515; AAB04134.1; -.
DR   EMBL; U58514; AAB04133.1; -.
DR   EMBL; BC000835; AAC50702.1; ALT_INIT.
DR   MIM; 601525; -.
DR   InterPro; IPR002557; Chitinase_2.
DR   InterPro; IPR001223; Glyco_hydro_18.
DR   Pfam; PF00704; Glyco_hydro_18; 1.
DR   PROSITE; PS01095; CHITINASE_18; FALSE_NEG.
KW   Hydrolase; Signal; Alternative splicing.
FT   SIGNAL       1     20       POTENTIAL.
FT   CHAIN       21    390       CHITINASE 3-LIKE PROTEIN 2.
FT   CARBOHYD    85     85       N-LINKED (GLCNAC...) (POTENTIAL).
FT   VARSPLIC     1     12       MLAWSQLCSLW -> MSSWMLLSFLVAGITGVHSHIIT
FT                                (IN ISOFORM 2).
FT                                ATIWRSRTRPGSTDFSPHPSI (IN ISOFORM 2).
SQ   SEQUENCE   390 AA;  [illegible] MW;  97B99A2F4AA15677 CRC64;


  Query Match             81.8%;  Score 45;  DB 1;  Length 390;
  Best Local Similarity   27.3%;  Pred. No. 2.7;
  Matches    9;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 FGRSFTLAS 9
             :|:|:|||
Db     260 YGRSFTLSA 268


RESULT   5
CHI1_MANSE
ID   CHI1_MANSE     STANDARD;      PRT:   554 AA.
AC   P29030;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   15-FEB-1995 (Rel. 31, Last annotation update)
DE   Endochitinase precursor (EC 3.2.1.14).
```

```
OS   Manduca sexta (Tobacco hawkmoth) (Tobacco hornworm).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
OC   Sphingoidea; Sphingidae; Sphinginae; Manduca.
OX   NCBI_TaxID=7130;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=93357793; PubMed=8395525;
RA   Kramer K.J., Corpuz L., Choi H.K., Muthukrishnan S.;
RT   "Sequence of a cDNA and expression of the gene encoding epidermal and
RT   gut chitinases of Manduca sexta.";
RL   Insect Biochem. Mol. Biol. 23:691-701(1993).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=97215580; PubMed=9061927;
RA   [illegible] H., Kramer K.J., Muthukrishnan S.;
RT   "Isolation and characterization of a genomic clone for the gene of an
RT   insect molting enzyme, chitinase.";
RL   Insect Biochem. Mol. Biol. 27:47-47(1997).
CC   -!- FUNCTION: DIGESTS CHITIN IN THE EXOSKELETON DURING THE MOLTING
CC       PROCESS.
CC   -!- CATALYTIC ACTIVITY: Hydrolysis of the 1,4-beta linkages of N-
CC       acetyl-D-glucosamine polymers of chitin.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- TISSUE SPECIFICITY: EPIDERMIS AND GUT.
CC   -!- DEVELOPMENTAL STAGE: HIGH LEVELS SEEN IN THE EPIDERMIS DURING
CC       MOLTING AND LARVAL-PUPAL AND PUPAL-ADULT ECDYSIS. LOW LEVELS
CC       ARE SEEN IN FEEDING LARVAE AND ADULTS. A RAPID RISE IN A
CC       RAPID DECLINE IS SEEN IN THE GUT IS DETECTED ON DAY 1 WITH LOW
CC       LEVELS SEEN ON DAYS 0, 7 AND 8.
CC   -!- SIMILARITY: BELONGS TO CHITINASE CLASS II (FAMILY 18 OF GLYCOSYL
CC       HYDROLASES).
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; U02270; AAC04924.1; -.
DR   EMBL; L49234; AAB53052.1; -.
DR   InterPro; IPR002557; Chitin_binding.
DR   InterPro; IPR001579; Chitinase_2.
DR   InterPro; IPR001223; Glyco_hydro_18.
DR   Pfam; PF01607; Chitin_bind_2; 1.
DR   Pfam; PF00704; Glyco_hydro_18; 1.
DR   SMART; SM00494; ChtBD2; 1.
DR   PROSITE; PS01095; CHITINASE_18; 1.
KW   Hydrolase; Glycosidase; Chitin degradation; Signal; Glycoprotein.
FT   SIGNAL       1     19       POTENTIAL.
FT   CHAIN       20    554       ENDOCHITINASE.
FT   DOMAIN     396    453       CBD/THR RICH.
FT   ACT_SITE   146    146       PROTON DONOR (BY SIMILARITY).
FT   CARBOHYD    85     85       N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD   303    303       N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD   407    407       N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD   545    545       N-LINKED (GLCNAC...) (POTENTIAL).
SQ   SEQUENCE   554 AA;  62203 MW;  3989D756C96CD490 CRC64;


  Query Match             79.5%;  Score 35;  DB 1;  Length 554;
  Best Local Similarity   66.7%;  Pred. No. 6.5;
  Matches    6;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 FGRSFTLAS 9
             :|:|:|||:
Db     270 YGRSFTLSA 278


RESULT   6
BCHX_RHOCA
```

```
ID   BCHX_RHOCA     STANDARD;      PRT;   333 AA.
AC   P26177;
DT   01-MAY-1992 (Rel. 22, Created)
DT   01-MAY-1992 (Rel. 22, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Chlorophyllide reductase 45.5 kDa chain (EC 1.3.1. ) (Chl ri
DE   reductase).
GN   BCHX.
OS   Rhodobacter capsulatus (Rhodopseudomonas capsulata).
OC   Bacteria; Proteobacteria; alpha subdivision; Rhodobacter group;
OC   Rhodobacter.
OX   NCBI_TaxID=1061;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=St. Louis;
RX   MEDLINE=93224464; PubMed=8468299;
RA   Burke D.H., Alberti M., Hearst J.E.;
RT   "The Rhodobacter capsulatus chlorin reductase-encoding locus, bchA,
RT   consists of three genes, bchX, bchY, and bchZ.";
RL   J. Bacteriol. 175:2407-2413(1993).
CC   -!- FUNCTION: CONVERTS CHLOROPHYLLS (CHL) INTO BACTERIOCHLOROPHYLLS
CC       (BCHL) BY REDUCING RING B OF THE TETRAPYRROLE.
CC   -!- PATHWAY: BACTERIOCHLOROPHYLL BIOSYNTHESIS.
CC   -!- SIMILARITY: BELONGS TO THE NIFH / BCHB / CHLL FAMILY.
CC   -----------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----------------------------------------------------------------
DR   EMBL; Z11165; CAA77548.1; -.
DR   PIR; S17828; S17828.
DR   PIR; B49850; B49850.
DR   HSSP; P00456; 1CP2.
DR   InterPro; IPR000392; Nitrogenase1.
DR   Pfam; PF00142; fer4_NifH; 1.
DR   PRINTS; PR00090; NITROGNASE1.
DR   PROSITE; PS00692; NIFH_FRXC_2; 1.
DR   PROSITE; PS00746; NIFH_FRXC_1; 1.
KW   Oxidoreductase; Photosynthesis; Bacteriochlorophyll biosynthesis;
KW   ATP-binding; Iron-sulfur; 4Fe-4S.
FT   NP_BIND      42     49       ATP (POTENTIAL).
FT   METAL       130    130       IRON-SULFUR (4FE-4S) (BY SIMILARITY).
FT   METAL       165    165       IRON-SULFUR (4FE-4S) (BY SIMILARITY).
SQ   SEQUENCE   333 AA;  35573 MW;  CD697D641C8943EE CRC64;

  Query Match             72.7%;  Score 32;  DB 1;  Length 333;
  Best Local Similarity   75.0%;  Pred. No. 17;
  Matches    6;  Conservative    2;  Mismatches    0;  Indels    2;  Gaps    1;

Qy       2 GRSFTLAS 9
           | | | |:
Db      47 GKSFILAN 54


RESULT   7
BCHA_RHOSH
ID   BCHX_RHOSH     STANDARD;      PRT;   333 AA.
AC   Q02431; Q9RFC2;
DT   01-JUL-1993 (Rel. 29, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Chlorophyllide reductase 45.5 kDa chain (EC 1.3.1. ) (Chl ri
DE   reductase).
GN   BCHX.
OS   Rhodobacter sphaeroides (Rhodopseudomonas sphaeroides).
OC   Bacteria; Proteobacteria; alpha subdivision; Rhodobacter group;
OC   Rhodobacter.
OX   NCBI_TaxID=1063;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 17023 / 2.4.1 / NCIB 8253 / DSM 158;
RX   MEDLINE=93173095; PubMed=8437569;
RA   McGlynn P., Hunter C.N.;
RT   "Genetic analysis of the bchC and bchA genes of Rb. sphaeroides.";
RL   Mol. Gen. Genet. 236:227-234(1993).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 17023 / 2.4.1 / NCIB 8253 / DSM 158;
RX   MEDLINE=20110911; PubMed=11148476;
RA   Choudhary M., Kaplan S.;
RT   "DNA sequence analysis of the pbeta operon from Rhodobacter
RT   sphaeroides 2.4.1.";
RL   Nucleic Acids Res. 28:862-867(2000).
CC   -!- FUNCTION: CONVERTS CHL B & PHYLL (CHL) INTO BACTERIO CHL B & PHYLL
CC       (BCHL) BY REDUCING RING B OF THE TETRAPYRROLE.
CC   -!- PATHWAY: BACTERIOCHLOROPHYLL BIOSYNTHESIS.
CC   -!- SIMILARITY: BELONGS TO THE NIFH / BCHB / CHLL FAMILY.
CC   -----------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----------------------------------------------------------------
DR   EMBL; AJ010302; CAB38747.1; -.
DR   EMBL; AF257...; AAF24297.1; -.
DR   PIR; S78270; S78270.
DR   PIR; S30915; S30915.
DR   HSSP; P00459; 1N2C.
DR   InterPro; IPR000392; Nitrogenase1.
DR   Pfam; PF00142; fer4_NifH; 1.
DR   PRINTS; PR00090; NITROGNASE1.
DR   PROSITE; PS00692; NIFH_FRXC_2; 1.
DR   PROSITE; PS00746; NIFH_FRXC_1; 1.
KW   Oxidoreductase; Photosynthesis; Bacteriochlorophyll biosynthesis;
KW   ATP-binding; Iron-sulfur; 4Fe-4S.
FT   NP_BIND      42     49       ATP (BY SIMILARITY).
FT   METAL       130    130       IRON-SULFUR (4FE-4S) (BY SIMILARITY).
FT   METAL       165    165       IRON-SULFUR (4FE-4S) (BY SIMILARITY).
FT   CONFLICT    127    127       A -> G (IN REF. 1).
SQ   SEQUENCE   333 AA;  35573 MW;  CD697D641C8943EE CRC64;


  Query Match             72.7%;  Score 32;  DB 1;  Length 333;
  Best Local Similarity   75.0%;  Pred. No. 17;
  Matches    6;  Conservative    2;  Mismatches    0;  Indels    2;  Gaps    1;

Qy       2 GRSFTLAS 9
           | | | |:
Db      47 GKSFILAN 54


RESULT   8
TKT_CAEEL
ID   TKT_CAEEL      STANDARD;      PRT;   434 AA.
AC   Q27371;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   15-JUL-1998 (Rel. 36, Last annotation update)
DE   Troponin I.
GN   MUP-2 OR T22E5.5.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditina;
OC   Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BRISTOL N.; TISSUE=Muscle;
```

```
RX   MEDLINE; 96[illegible].  PubMed 8[illegible].
RA   Myers E.W., Gol[illegible]A., Allen J.S., Bonner E.A., Bonnett J.;
RT   "Developmental genetic analysis of troponin I mutations in striated
RT   and nonstriated muscle cells of Caenorhabditis elegans.";
RL   J. Cell Biol. 1[illegible]:[illegible] 1077(1996).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN BRISTOL N2;
RA   Minx P.;
RL   Submitted (JAN-1996) to the EMBL/GenBank/DDBJ databases.
CC   !  FUNCTION: TR-TROPIN I IS THE TROPOMYOSIN BINDING SUBUNIT OF
CC       TROPONIN, THE THIN FILAMENT REGULATORY COMPLEX WHICH CONFERS
CC       CALCIUM SENSITIVITY TO STRIATED MUSCLE ACTOMYOSIN ATPASE ACTIVITY
CC       (BY SIMILARITY).
CC   !  SIMILARITY: TO VERTEBRATE SKELETAL AND CARDIAC TNI.
CC   -----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----------------------------------------------------------------------
DR   EMBL; U14775; AAA80[illegible].1; -.
DR   EMBL; U15[illegible]; AAA80[illegible].1; -.
DR   WormPep; T22E5.[illegible]; CE04964.
DR   InterPro; IPR00[illegible]; Troponin.
DR   Pfam; PF00992; Troponin; 1.
FT   DOMAIN       11      23       POLY-GLU.
FT   DOMAIN       [illegible]      78       POLY-GLU.
SQ   SEQUENCE   405 AA;  47341 MW;  2[illegible]A[illegible]EEB641[illegible]5B CRC64;

  Query Match             72.7%;  Score 32;  DB 1;  Length 405;
  Best Local Similarity   75.0%;  Pred. No. 20;
  Matches    6;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2 GRSFTLAS 9
           |||:|:|
Db     124 GRNFIVAS 141


RESULT   9
ZNF2_MESSA
ID   ZNF2_MESSA     STANDARD;      PRT:   415 AA.
AC   P40621;
DT   01-OCT-1994 (Rel. 31, Created)
DT   01-OCT-1994 (Rel. 31, Last sequence update)
DT   01-OCT-1994 (Rel. 31, Last annotation update)
DE   Zinc finger protein ZFMSAI2A.
OS   Micropterus salmoides (Largemouth bass).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC   Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Percoidei;
OC   Centrarchidae; Micropterus.
OX   NCBI_TaxID=2[illegible];
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Ovary;
RX   MEDLINE; 9[illegible]; PubMed 8[illegible];
RA   [illegible]M.S., Smith T.I.J., Lewis C.S.;
RT   "Sequence similarities of a TFIIIA-type zinc finger protein from fish
RT   ovarian tissue.";
RL   Nucleic Acids Res. 21:[illegible] 4[illegible]_4[illegible](1993).
CC   !  SUBCELLULAR LOCATION: Nuclear (Probable).
CC   -----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----------------------------------------------------------------------
DR   EMBL; Z22756; CAA80448.1; -.
DR   PIR; S44140; S44140.
DR   HSSP; P08047; 1SP2.
DR   InterPro; IPR000822; Znf-C2H2.
DR   Pfam; PF00096; zf-C2H2; 12.
DR   PRINTS; PR00048; ZINCFINGER.
DR   SMART; SM00355; ZnF_C2H2; 12.
DR   PROSITE; PS00028; ZINC_FINGER_C2H2_1; 11.
DR   PROSITE; PS50157; ZINC_FINGER_C2H2_2; 12.
KW   Zinc finger; Metal binding; DNA binding; Repeat; Nuclear protein.
FT   ZN_FING      76     100       C2H2-TYPE.
FT   ZN_FING      106    129       C2H2-TYPE.
FT   ZN_FING      134    156       C2H2-TYPE.
FT   ZN_FING      161    183       C2H2-TYPE.
FT   ZN_FING      189    211       C2H2-TYPE.
FT   ZN_FING      217    239       C2H2-TYPE.
FT   ZN_FING      245    267       C2H2-TYPE.
FT   ZN_FING      273    295       C2H2-TYPE.
FT   ZN_FING      301    323       C2H2-TYPE.
FT   ZN_FING      329    351       C2H2-TYPE.
FT   ZN_FING      357    379       C2H2-TYPE.
FT   ZN_FING      385    407       C2H2-TYPE.
SQ   SEQUENCE   415 AA;  47237 MW;  8A2D2B90C4027B08 CRC64;

  Query Match             72.7%;  Score 32;  DB 1;  Length 415;
  Best Local Similarity   75.0%;  Pred. No. 21;
  Matches    6;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2 GRSFTLAS 9
           |||:|:|
Db     134 GRGFTMAS 142


RESULT  10
FAT1_SCHPO
ID   FAT1_SCHPO     STANDARD;      PRT:  1385 AA.
AC   O14236; Q9UTJ0;
DT   15-JUL-1998 (Rel. 36, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Actin interacting protein 3 homolog.
GN   FAT1 OR SPAC15A10.15 OR SPAC15E1.01.
OS   Schizosaccharomyces pombe (Fission yeast).
OC   Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC   Schizosaccharomycetales; Schizosaccharomycetaceae;
OC   Schizosaccharomyces.
OX   NCBI_TaxID=4896;
RN   [1]
RP   SEQUENCE OF 1-1033 FROM N.A.
RC   STRAIN 972;
RA   Murphy L., Harris D., Wood V., Barrell B.G., Rajandream M.A.;
RL   Submitted (OCT-2001) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE OF 1023-1385 FROM N.A.
RC   STRAIN 972;
RA   Murphy L., Harris D., Barrell B.G., Rajandream M.A., Lyne M.H.;
RL   Submitted (APR-2002) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   IDENTIFICATION, AND GENE NAME.
RX   MEDLINE; 20144585; PubMed 10879021;
RA   Jin H., Ambera D.C.;
RT   "The secretory pathway mediates localization of the cell growth
RT   regulator Aip3p/Bud6p.";
RL   Mol. Biol. Cell 11:647-661(2000).
CC   !  FUNCTION: INVOLVED IN THE ORGANIZATION AND/OR FUNCTION OF THE
CC       ACTIN CYTOSKELETON.
CC   !  SIMILARITY: TO YEAST BUD6.
CC   -----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
```

CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

DR EMBL; Z97108; CAB01112.1; -.
DR EMBL; AL109770; CAB42420.1; -.
KW Coiled coil; Cytoskeleton.
FT DOMAIN 1009 1096 COILED COIL (POTENTIAL).
SQ SEQUENCE 1385 AA; 154425 MW; AB83B4B07B4FF7E7 CRC64;

Query Match 72.7%; Score 22; DB 1; Length 1385;
Best Local Similarity 66.7%; Pred. No. 78;
Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 FGKSFILAS 9
|||:|:|
Db 1009 FGRGFVLSS 1017

RESULT 11
ZN36_HUMAN
ID ZN36_HUMAN STANDARD; PRT; 325 AA.
AC P17029; P52745;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Zinc finger protein 36 (Zinc finger protein KOX18) (Fragment).
GN ZNF36 OR KOX18 OR ZNF39.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Fibroblast;
RX MEDLINE=90344??; PubMed=2557960;
RA Tommerup N., Vissing H.;
RT "Isolation and fine mapping of 16 novel human zinc finger-encoding
RT cDNAs: Identify putative candidate genes for developmental and
RT malignant disorders.";
RL Genomics 7:265-274(1990).
RN [2]
RP SEQUENCE OF 197-222 FROM N.A.
RC TISSUE=Lymphoid;
RX MEDLINE=91145439; PubMed=2288909;
RA Thiesen H.-J.;
RT "Multiple genes encoding zinc finger domains are expressed in human T
RT cells.";
RL New Biol. 2:363-374(1990).
CC -!- FUNCTION: MAY BE INVOLVED IN TRANSCRIPTIONAL REGULATION.
CC -!- SUBCELLULAR LOCATION: Nuclear (Potential).
CC -!- SIMILARITY: BELONGS TO THE KRUEPPEL FAMILY OF C2H2 TYPE ZINC
CC FINGER PROTEINS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

DR EMBL; D00849; AAC50761.1; -.
DR EMBL; X12349; CAA30575.1; -.
DR PIR; S14406; S14406.
DR HSSP; P25490; 1ZNM.
DR MIM; 600126; -.
DR InterPro; IPR000822; Znf-C2H2.
DR Pfam; PF00096; zf-C2H2; 6.
DR PRINTS; PR00048; ZINCFINGER.
DR SMART; SM00355; ZnF_C2H2; 6.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 7.
DR PROSITE; PS50157; ZINC_FINGER_C2H2_2; 6.
KW Transcription regulation; DNA-binding; Zinc-finger; Metal-binding;
KW Nuclear protein; Repeat.
FT NON_TER 1 1
FT DOMAIN 139 301 ZINC FINGERS.
FT ZN_FING 139 161 C2H2 TYPE.
FT ZN_FING 167 189 C2H2 TYPE.
FT ZN_FING 195 217 C2H2 TYPE.
FT ZN_FING 223 245 C2H2 TYPE.
FT ZN_FING 251 273 C2H2 TYPE.
FT ZN_FING 279 301 C2H2 TYPE.
SQ SEQUENCE 325 AA; 36643 MW; 2D7B7C4B62B65A0A CRC64;

Query Match 2.5%; Score 8; DB 1; Length 325;
Best Local Similarity 75.0%; Pred. No. 2.7;
Matches 6; Conservative 2; Mismatches 0; Indels 1; Gaps 1;

Qy 2 GRSFILAS 9
|||:|:|
Db 117 GRSFSLSS 124

RESULT 12
TENS_CHICK
ID TENS_CHICK STANDARD; PRT; 1744 AA.
AC Q04205; Q91007; Q92611;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Tensin.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Heart;
RX MEDLINE=94356087; PubMed=8071458;
RA Lo S.H., An Q., Bao S., Wong W.K., Liu Y., Dunney P.A., Barwick T.H.,
RA Chen L.B.;
RT "Molecular cloning of chick cardiac muscle tensin. Full coding cDNA
RT sequence, expression, and characterization.";
RL J. Biol. Chem. 269:22810-22820(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Heart;
RX MEDLINE=95204540; PubMed=7896874;
RA Chuang Z.Z., Lin C.C., Lin S.;
RT "Molecular cloning, expression, and mapping of the 5'end actin-
RT actin cappind domain of chicken cardiac tensin.";
RL J. Cell Biol. 128:1095-1109(1995).
RN [3]
RP SEQUENCE FROM N.A.
RA Chen L.B.;
RL Submitted (XXX-1994) to the EMBL/GenBank/DDBJ databases.
RN [4]
RP SEQUENCE OF 1459-1744 FROM N.A.
RC TISSUE=Embryonic chondrocytes and Embryonic heart;
RX MEDLINE=94039118; PubMed=8262051;
RA van de Werken R., Gennari M., Tavella S., Bet P., Molina F.,
RA Lin S., Cancedda R., Castagnola P.;
RT "Modulation of tensin and vinculin expression in chick embryo
RT developing cartilage and cultured differentiating chondrocytes.";
RL Eur. J. Biochem. 217:781-790(1993).
RN [5]
RP SH2 DOMAIN.
RX MEDLINE=91220073; PubMed=1709717;
RA Davis S., Lu M.L., Lo S.H., Butler J.A., Druker B.J.,
RA Roberts T.M., An Q., Chen L.B.;

```
RT   "Phosphorylation of the SH2 domain of the actin-binding protein tensin.";
RL   Science 252:71-715 (1991).
CC   -!- FUNCTION: MAY BE INVOLVED IN CELL MIGRATION, CANCERGENESIS AND
CC       AND IN LINKING SIGNAL TRANSDUCTION PATHWAYS TO THE CYTOSKELETON.
CC       BINDS TO ACTINS AND PHOSPHORYLATES PROTEINS ON SER TRANSFORMED
CC       CELLS, MAY BIND ACTIN WITH CAPPING AND BUNDLING PROPERTIES.
CC   -!- SUBCELLULAR LOCATION: LOCALIZED TO ADHERENS JUNCTIONS.
CC   -!- TISSUE SPECIFICITY: BRAIN, PLACENTA, LUNG AND SKELETAL MUSCLE.
CC   -!- PTM: TYROSINE PHOSPHORYLATED.
CC   -!- SIMILARITY: CONTAINS 1 TENSIN DOMAIN.
CC   -!- SIMILARITY: CONTAINS 1 SH2 DOMAIN.
CC   -----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----------------------------------------------------------------------
DR   EMBL; M96822; AAA96134.1; -.
DR   EMBL; L07061; AAA73610.1; ALT_INIT.
DR   EMBL; Z18520; CAA79195.1; ALT_INIT.
DR   EMBL; M74105; AAA60027.1; -.
DR   EMBL; X66280; CAA46962.1; -.
DR   HSSP; P62753; 1BLI.
DR   Prosite; PF00069; PKD; 1.
DR   InterPro; IPR000980; SH2.
DR   Pfam; PF00017; SH2; 1.
DR   SMART; SM00252; PTB; 1.
DR   SMART; SM00326; SH3; 1.
DR   SMART; SM00252; SH2; 1.
DR   PROSITE; PS50001; SH2; 1.
KW   Actin-binding; Cytoskeleton; SH2 domain; Phosphorylation.
FT   DOMAIN       96    312       TENSIN.
FT   DOMAIN      147   1961       SH2.
FT   CONFLICT     49     49       G -> R (IN REF. 2).
FT   CONFLICT     61     61       M -> I (IN REF. 2).
FT   CONFLICT     83     83       - -> IK (IN REF. 3).
FT   CONFLICT    404    404       A -> I (IN REF. 2).
FT   CONFLICT    452    452       I -> A (IN REF. 2).
FT   CONFLICT    508    509       DV -> EL (IN REF. 2).
FT   CONFLICT    522    522       A -> E (IN REF. 2).
FT   CONFLICT    604    604       R -> A (IN REF. 2).
FT   CONFLICT    666    666       A -> I (IN REF. 2).
FT   CONFLICT    875    875       I -> A (IN REF. 2).
FT   CONFLICT    909    909       I -> L (IN REF. 2).
FT   CONFLICT   1102   1103       MISSING (IN REF. 3).
FT   CONFLICT   1238   1238       S -> A (IN REF. 2).
FT   CONFLICT   1389   1389       E -> D (IN REF. 2).
FT   CONFLICT   1711   1711       D -> E (IN REF. 2).
SQ   SEQUENCE   1744 AA;  197,134 MW;  F3D9B...211A5F5A CRC64;


  Query Match             57.2%;  Score 41;  DB 1;  Length 1744;
  Best Local Similarity   86.7%;  Pred. No. 1.5e-02;
  Matches    13;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 FGRSFTLAS 9
             ||| ||| |
Db    1750 FGNSFIVDS 1758


RESULT  13
R14A_YEAST
ID   R14A_YEAST     STANDARD;      PRT:   199 AA.
AC   P12709;
DT   01-NOV-1990 (Rel. 15, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   15-DEC-1998 (Rel. 37, Last annotation update)
DE   60S ribosomal protein L14-A.
GN   RPL14A OR RPL14 OR YML073C.
OS   Saccharomyces cerevisiae (Baker's yeast).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX   NCBI_TaxID=4932;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Wimmer K., Wolfer H., Wolfer E., Schafer M.;
RL   Submitted (JUL-1990) to the EMBL/GenBank/DDBJ databases.
CC   -!- MISCELLANEOUS: THERE ARE TWO GENES FOR L14 IN YEAST.
CC   -!- SIMILARITY: BELONGS TO THE L14E FAMILY OF RIBOSOMAL PROTEINS.
CC   -----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----------------------------------------------------------------------
DR   EMBL; Z47071; CAA87356.1; -.
DR   SGD; S0004750; RPL14B.
DR   InterPro; IPR001380; Ribosomal_L14E.
DR   Pfam; PF01294; Ribosomal_L14e; 1.
DR   ProDom; PD004443; Ribosomal_L14E; 1.
DR   PROSITE; PS01104; RIBOSOMAL_L14E; 1.
KW   Ribosomal protein; Multigene family.
SQ   SEQUENCE   199 AA;  22554 MW;  EA7C95F1F80345A CRC64;


  Query Match             68.2%;  Score 49;  DB 1;  Length 199;
  Best Local Similarity   85.7%;  Pred. No. 26;
  Matches    6;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       2 GRSFTLA 8
             ||| |||
Db     72 GRGFTLA 78


RESULT  14
R14B_YEAST
ID   R14B_YEAST     STANDARD;      PRT:   199 AA.
AC   P14712;
DT   01-FEB-1995 (Rel. 31, Created)
DT   01-FEB-1995 (Rel. 31, Last sequence update)
DT   15-DEC-1998 (Rel. 37, Last annotation update)
DE   60S ribosomal protein L14-B.
GN   RPL14B OR RPL14 OR YMR142C OR YM9675.11C.
OS   Saccharomyces cerevisiae (Baker's yeast).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX   NCBI_TaxID=4932;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN S288C / AB972;
RA   Badcock K., Churcher C., Barrell B.G., Rajandream M.A., Walsh S.V.;
RL   Submitted (DEC-1994) to the EMBL/GenBank/DDBJ databases.
CC   -!- MISCELLANEOUS: THERE ARE TWO GENES FOR L14 IN YEAST.
CC   -!- SIMILARITY: BELONGS TO THE L14E FAMILY OF RIBOSOMAL PROTEINS.
CC   -----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----------------------------------------------------------------------
DR   EMBL; Z47071; CAA87356.1; -.
DR   SGD; S0004750; RPL14B.
DR   InterPro; IPR001380; Ribosomal_L14E.
DR   Pfam; PF01294; Ribosomal_L14e; 1.
DR   ProDom; PD004443; Ribosomal_L14E; 1.
DR   PROSITE; PS01104; RIBOSOMAL_L14E; 1.
```

```
KW   ribosomal protein; MG1 intene family.
SQ   SEQUENCE   199 AA;  22525 MW;  BEA4B42CEE76FF23 CRC64;


  Query Match             68.2%;  Score 30;  DB 1;  Length 199;
  Best Local Similarity   85.7%;  Pred. No. 26;
  Matches    6;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       2 GRSFILA 8
           || ||||
Db      72 GRGFILA 79


RESULT  15
RL13_CANAL
ID   RL13_CANAL     STANDARD;      PRT:   202 AA.
AC   O59931;
DT   15-DEC-1998 (Rel. 37, Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   60S ribosomal protein L13.
GN   RPL13 OR CA4901a.32.
OS   Candida albicans (Yeast).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; mitosporic Saccharomycetales; Candida.
OX   NCBI TaxID=5476;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN SC5314;
RA   Water X., Andrews B.L., Polta J.M., Sundstrom P.;
RT   "Sequence of the Candida albicans cDNA encoding ribosomal protein
RT   L13.";
RL   Submitted (DEC-1998) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN 1161;
RA   Taylor K., Barris L., Barrell B.G., Rajandream M.A.;
RL   Submitted (NOV-1998) to the EMBL/GenBank/DDBJ databases.
CC   !  SIMILARITY: BELONGS TO THE L13E FAMILY OF RIBOSOMAL PROTEINS.
CC   -----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----------------------------------------------------------------------
DR   EMBL; AF057972; AAD09956.1;  -.
DR   EMBL; D86864; AAB09226.1;  -.
DR   EMBL; AL033497; CAA21966.1;  -.
DR   InterPro; IPR001380; Ribosomal_L13E.
DR   Pfam; PF01294; Ribosomal_L13e; 1.
DR   Prodom; PD000444; Ribosomal_L13E; 1.
DR   PROSITE; PS01164; RIBOSOMAL_L13E; 1.
KW   Ribosomal protein.
SQ   SEQUENCE   202 AA;  23053 MW;  97223A2714BA784E CRC64;


  Query Match             68.2%;  Score 30;  DB 1;  Length 202;
  Best Local Similarity   85.7%;  Pred. No. 26;
  Matches    6;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       2 GRSFILA 8
           || ||||
Db      75 GRGFILA 82


Search completed: August 26, 2002, 06:29:00
Job time: 929 sec
```

Smith-Waterman search, version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

Comparison matrix: BLOSUM62
Gap_open penalty: 10
Gap_extension penalty: 0.5

Perform protein - protein search, using sw model

Run on: August 26, 2002, 06:12:54 ; Search time 13.16 Seconds
(without alignments)
11,959 Million cell updates/sec

Title: US-09-744-282-1
Perfect score: 44
Sequence: 1 EGRSFILAS 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562,222 seqs, 172,001,929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_19:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_other:*
8: sp_rodent:*
9: sp_todent:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriap:*
17: sp_archeap:*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	44	100.0	383	4	Q96HI7	Q96HI7 homo sapien
2	44	100.0	383	6	Q29411	Q29411 sus scrofa
3	43	97.7	382	6	Q9N4K5	Q9N4K5 oryctolag
4	42	95.5	382	11	Q9WIV1	Q9WIV1 rattus nor
5	42	95.5	383	12	Q99ZB4	Q99ZB4 gga gallus
6	42	95.5	387	4	Q9HBV8	Q9HBV8 homo sapien
7	42	95.5	399	12	Q9DCL1	Q9DCL1 gga gallus
8	42	95.5	469	4	Q7Z2Z1	Q7Z2Z1 homo sapien
9	38	86.4	466	5	Q9W2M7	Q9W2M7 drosophila
10	38	86.4	529	5	Q1HC79	Q1HC79 anopheles g
11	38	86.4	958	5	P91741	P91741 hyphantria
12	38	86.4	966	5	Q9VZV2	Q9VZV2 drosophila
13	38	86.4	1018	5	Q9V6M0	Q9V6M0 drosophila
14	38	86.4	1276	5	Q9V1Z4	Q9V1Z4 drosophila
15	38	86.4	491	4	Q9HE47	Q9HE47 homo sapien
16	38	86.4	1373	14	Q9X2S7	Q9X2S7 liermodera
17	38	86.4	543	5	Q9GV05	Q9GV05 bombyx mori
18	38	86.4	543	5	Q9GR93	Q9GR93 bombyx mori
19	38	86.4	544	5	Q9GQC4	Q9GQC4 bombyx mori
20	38	86.4	552	5	Q9GV44	Q9GV44 spodoptera
21	38	86.4	565	5	Q9GDG9	Q9GDG9 bombyx mori
22	38	86.4	566	5	P98710	P98710 bombyx mori
23	38	86.4	765	10	Q9SHL3	Q9SHL3 arabidopsis
24	34	77.3	414	16	Q9KAQ3	Q9KAQ3 bacillus ha
25	34	77.3	1036	3	Q9HBQ9	Q9HBQ9 emericella
26	34	77.3	1080	5	Q9BLI6	Q9BLI6 bombyx mori
27	33	75.0	122	2	Q9EX60	Q9EX60 streptomyce
28	33	75.0	249	2	Q9L4B3	Q9L4B3 rhizobium l
29	33	75.0	267	16	Q9PDV7	Q9PDV7 xylella fas
30	33	75.0	405	5	O97403	O97403 phaedon co
31	33	75.0	644	4	Q9Y2K1	Q9Y2K1 homo sapien
32	33	75.0	684	17	Q29784	Q29784 archaeal-b
33	33	75.0	713	11	Q91VL9	Q91VL9 mus musculu
34	33	75.0	1635	5	O17412	O17412 aedes aegyp
35	32	72.7	114	17	O59326	O59326 pyrococcus
36	32	72.7	269	12	Q91U91	Q91U91 grapevine l
37	32	72.7	294	5	Q9N4K5	Q9N4K5 caenorhabdi
38	32	72.7	384	2	Q9JPC1	Q9JPC1 rhodocyclus
39	32	72.7	403	17	Q9HLJ3	Q9HLJ3 thermoplasm
40	32	72.7	488	5	Q24737	Q24737 chelonus sp
41	32	72.7	574	5	O17411	O17411 aedes aegyp
42	32	72.7	592	2	Q92N80	Q92N80 sinorhizobi
43	32	72.7	595	5	Q9VFK3	Q9VFK3 drosophila
44	32	72.7	741	5	Q19535	Q19535 caenorhabdi
45	32	72.7	744	5	Q95QK5	Q95QK5 caenorhabdi

ALIGNMENTS

RESULT 1
Q96HI7
ID Q96HI7 PRELIMINARY; PRT; 383 AA.
AC Q96HI7;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE SIMILAR TO CHITINASE 3 LIKE 1 (CARTILAGE GLYCOPROTEIN 39).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BRAIN, AND GLIOBLASTOMA WITH EGFR AMPLIFICATION;
RA Strausberg R.;
RL Submitted (MAY 2001) to the EMBL/GenBank/DDBJ databases.
DR EMBL; BC008568; AAH08568.1; -.
SQ SEQUENCE 383 AA; 42425 MW; 6C9EF183BDC7C2D1 CRC64;

Query Match 100.0%; Score 44; DB 4; Length 383;
Best Local Similarity 100.0%; Pred. No. 0.28;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EGRSFILAS 9
|||||||||
Db 241 EGRSFILAS 249

RESULT 2
Q29411
ID Q29411 PRELIMINARY; PRT; 383 AA.
AC Q29411;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE 38 KDA HEPARIN BINDING GLYCOPROTEIN.
OS Sus scrofa (Pig).

```
OS   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX   NCBI_TaxID=9823;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=SMOOTH MUSCLE;
RX   MEDLINE=95286589; PubMed=7768902;
RA   Shackelton L.M., Mann D.M., Millis A.J.;
RT   "Identification of a 48 kDa heparin-binding glycoprotein (apo48) in
RT   differentiating vascular smooth muscle cells as a member of a group of
RT   proteins associated with tissue remodeling.";
RL   J. Biol. Chem. 270:13076-13083(1995).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=SMOOTH MUSCLE;
RA   Shackelton L.M., Mann D.M., Millis A.J.T.;
RL   Submitted (JAN-1995) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U19900; AAA86482.1; -.
DR   EMBL; Z47803; CAA87764.1; -.
DR   InterPro; IPR001223; Glyco_hydro_18.
DR   Pfam; PF00704; Glyco_hydro_18; 1.
KW   Heparin binding.
SQ   SEQUENCE   388 AA;  42443 MW;  4D1049E4991DBDC4 CRC64;


  Query Match             100.0%;  Score 44;  DB 6;  Length 388;
  Best Local Similarity   100.0%;  Pred. No. 0.28;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 FGRSFTLAS 9
           |||||||||
Db     262 FGRSFTLAS 270


RESULT   3
Q18945
ID   Q18945      PRELIMINARY;      PRT;   332 AA.
AC   Q18945;
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 15, Last annotation update)
DE   CHITINASE LIKE PROTEIN 1 (FRAGMENT).
GN   CLP 1.
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC   Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Recklies A.D., White C.;
RT   "Expression of chitinase like protein 1 (CLP-1) in bovine
RT   chondrocytes.";
RL   Submitted (JUN-1997) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF011478; AAB64304.1; -.
DR   InterPro; IPR001223; Glyco_hydro_18.
DR   Pfam; PF00704; Glyco_hydro_18; 1.
FT   NON_TER       1       1
FT   NON_TER     332     332
SQ   SEQUENCE   332 AA;  37434 MW;  03F16323486408F CRC64;


  Query Match             93.2%;  Score 41;  DB 6;  Length 332;
  Best Local Similarity   88.9%;  Pred. No. 1.1;
  Matches    8;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 FGRSFTLAS 9
           ||||:||||
Db     247 FGRSYTLAS 255


RESULT   4
Q9WIV1
ID   Q9WIV1      PRELIMINARY;      PRT;   352 AA.
AC   Q9WIV1;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 15, Last annotation update)
DE   GLYCOPROTEIN 39 (FRAGMENT).
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=LEWIS;
RA   Wendling U., Boots A.M.H., van Eden W.;
RT   "Cloning of the rat homologue of Human Cartilage glycoprotein-39 a
RT   potential autoantigen in arthritis.";
RL   Submitted (AUG-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF062038; AAD22610.1; -.
DR   HSSP; P07254; 1C1N.
DR   InterPro; IPR000677; 2S_Globulin.
DR   InterPro; IPR001223; Glyco_hydro_18.
DR   Pfam; PF00704; Glyco_hydro_18; 1.
DR   PRINTS; PR00551; 2SGLOBULIN.
FT   NON_TER       1       1
FT   NON_TER     352     352
SQ   SEQUENCE   352 AA;  37651 MW;  C1B2F6616B0A0C6AC CRC64;


  Query Match             93.2%;  Score 41;  DB 11;  Length 352;
  Best Local Similarity   88.9%;  Pred. No. 1.1;
  Matches    8;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 FGRSFTLAS 9
           ||:||||||
Db     240 FGKSFTLAS 248


RESULT   5
Q99J84
ID   Q99J84      PRELIMINARY;      PRT;   391 AA.
AC   Q99J84;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-OCT-2001 (TrEMBLrel. 18, Last annotation update)
DE   SIMILAR TO CHITINASE 3 LIKE 1 (CARTILAGE GLYCOPROTEIN 39).
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Strausberg R.;
RL   Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Strausberg R.;
RL   Submitted (FEB-2001) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RA   Strausberg R.;
RL   Submitted (MAR-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC005611; AAH05611.1; -.
DR   EMBL; BC004790; AAH04790.1; -.
DR   EMBL; BC004774; AAH04774.1; -.
DR   InterPro; IPR001223; Glyco_hydro_18.
DR   Pfam; PF00704; Glyco_hydro_18; 1.
SQ   SEQUENCE   391 AA;  42979 MW;  EF65AA6AE4445C CRC64;


  Query Match             93.2%;  Score 41;  DB 11;  Length 391;
  Best Local Similarity   88.9%;  Pred. No. 1.2;
  Matches    8;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy       1 FGRSFILAS 9
           ^ ||:| :|||
Db     262 FGKSFILAS 270


RESULT   5
Q9H3V8
ID   Q9H3V8      PRELIMINARY;      PRT:   397 AA.
AC   Q9H3V8;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   CHITOTRIOSIDASE-LIKE PROTEIN.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20364067; PubMed=11217851;
RA   Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA   Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA   Aizawa K., Izawa M., Nishi F., Kiyosawa H., Kondo S., Yamanaka I.,
RA   Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA   Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA   Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA   Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA   Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA   Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA   Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.J.,
RA   Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA   Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA   Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA   Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA   Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.F.,
RA   Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA   Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA   Hayashizaki Y.;
RT   "Functional annotation of a full-length mouse cDNA collection.";
RL   Nature 409:685-690(2001).
CC   -!- SIMILARITY: BELONGS TO CHITINASE CLASS II (FAMILY 18 OF GLYCOSYL
CC       HYDROLASES).
DR   EMBL; AK009012; BAB26025.1; -.
DR   MGD; MGI:1919134; 2300002I19Rik.
DR   InterPro; IPR000677; ZS_Globulin.
DR   InterPro; IPR001579; Chitinase_2.
DR   InterPro; IPR001223; Glyco_hydro_18.
DR   Pfam; PF00704; Glyco_hydro_18; 1.
DR   PRINTS; PR00551; ZSGLOBULIN.
DR   PROSITE; PS01095; CHITINASE_18; 1.
KW   Glycosidase; Hydrolase.
SQ   SEQUENCE   396 AA;  43613 MW;  42BB674B61AF7128 CRC64;


  Query Match             83.2%;  Score 41;  DB 11;  Length 396;
  Best Local Similarity   88.9%;  Pred. No. 1.4;
  Matches    8;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 FGRSFILAS 9
           :|||||||:|
Db     267 YGRSFILAS 275


RESULT   6
Q13231
ID   Q13231      PRELIMINARY;      PRT:   466 AA.
AC   Q13231;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   CHITOTRIOSIDASE PRECURSOR.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=96064695; PubMed=7592832;
RA   Boot R.G., Renkema G.H., Strijland A., van Zonneveld A.J., Aerts J.M.F.;
RT   "Cloning of a cDNA encoding chitotriosidase, a human chitinase
RT   produced by macrophages.";
RL   J. Biol. Chem. 270:26252-26256(1995).
CC   -!- SIMILARITY: BELONGS TO CHITINASE CLASS II (FAMILY 18 OF GLYCOSYL
CC       HYDROLASES).
DR   EMBL; U29615; AAC50246.1; -.
DR   InterPro; IPR001579; Chitinase_2.
DR   InterPro; IPR002557; Chitin_bindint.
DR   InterPro; IPR001223; Glyco_hydro_18.
DR   Pfam; PF01607; Chitin_bind_2; 1.
DR   Pfam; PF00704; Glyco_hydro_18; 1.
DR   SMART; SM00494; ChtBD2; 1.
DR   PROSITE; PS01095; CHITINASE_18; 1.
KW   Glycosidase; Hydrolase; Signal.
FT   SIGNAL       1     21       POTENTIAL.
FT   CHAIN       22    466       CHITOTRIOSIDASE.
```

```
SQ   SEQUENCE   466 AA;  51681 MW;  B431201EB0F6E863 CRC64;


  Query Match             93.2%;  Score 41;  DB 4;  Length 466;
  Best Local Similarity   88.9%;  Pred. No. 1.5;
  Matches    8;  Conservative    1;  Mismatches    0;  Indels    1;  Gaps    0;

Qy       1 FGRSFTLAS 9
           :|:|:|||:
Db     267 YGRSFTLAS 275


RESULT   9
Q9W2M7
ID   Q9W2M7      PRELIMINARY;      PRT;   466 AA.
AC   Q9W2M7;
DT   01-MAY-2000 (TrEMBLrel. 14, Created)
DT   01-MAY-2000 (TrEMBLrel. 14, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   CG9657 PROTEIN.
GN   CG9657.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BERKELEY;
RX   MEDLINE=20076046; PubMed=10731132;
RA   Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA   Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA   George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA   Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA   Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA   Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA   Abril J.F., Agbayani A., An H., Andrews-Pfannkoch C., Baldwin D.,
RA   Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA   Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA   Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA   Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA   Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA   de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA   Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C.,
RA   Dunn P., Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA   Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA   Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA   Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA   Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA   Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA   Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA   Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA   Liu X., Mattei B., McIntosh T.C., Mcleod M.P., McPherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA   Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA   Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT   "The genome sequence of Drosophila melanogaster.";
RL   Science 287:2185-2195(2000).
CC   -!- SIMILARITY: BELONGS TO CHITINASE CLASS II (FAMILY 18; EC 3.2.1.14; SYN.
CC       HYDROLASES).
DR   EMBL; AE003452; AAF46663.1; -.
DR   HSSP; P07254; 1CTN.
DR   FlyBase; FBgn0034580; CG9657.
DR   InterPro; IPR000726; Chitinase_2.
DR   InterPro; IPR002557; Chitin_binding.
DR   InterPro; IPR000773; EXBD_like.
DR   InterPro; IPR000_223; Glyco_hydro_18.
DR   Pfam; PF01607; Chitin_bind_2; 1.
DR   Pfam; PF00704; Glyco_hydro_18; 1.
DR   SMART; SM00494; ChtBD2; 1.
DR   PROSITE; PS01095; CHITINASE_18; 1.
DR   PROSITE; PS00545; EXBP_PFIAGE_1; UNKNOWN 1.
KW   Glycosidase; Hydrolase.
SQ   SEQUENCE   466 AA;  51615 MW;  BEE41126E1106C0B CRC64;


  Query Match             89.4%;  Score 39;  DB 5;  Length 466;
  Best Local Similarity   77.8%;  Pred. No. 6.7;
  Matches    7;  Conservative    2;  Mismatches    0;  Indels    1;  Gaps

Qy       1 FGRSFTLAS 9
           :|||||||:
Db     264 YGRSFTLAI 272


RESULT   10
O44079
ID   O44079      PRELIMINARY;      PRT;   472 AA.
AC   O44079;
DT   01-JUN-1998 (TrEMBLrel. 09, Created)
DT   01-JUN-1998 (TrEMBLrel. 09, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   CHITINASE.
GN   AGCHI-1.
OS   Anopheles gambiae (African malaria mosquito).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicidae;
OC   Anopheles.
OX   NCBI_TaxID=7165;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=GUT;
RX   MEDLINE=98049056; PubMed=9405658;
RA   Shen Z., Jacobs-Lorena M.;
RT   "Characterization of a novel gut-specific chitinase gene from the
RT   human malaria vector Anopheles gambiae.";
RL   J. Biol. Chem. 272:28895-28900(1997).
CC   -!- SIMILARITY: BELONGS TO CHITINASE CLASS II (FAMILY 18; EC 3.2.1.14; SYN.
CC       HYDROLASES).
DR   EMBL; AF008575; AAB97764.1; -.
DR   InterPro; IPR001579; Chitinase_2.
DR   InterPro; IPR002557; Chitin_binding.
DR   InterPro; IPR001223; Glyco_hydro_18.
DR   Pfam; PF01607; Chitin_bind_2; 1.
DR   Pfam; PF00704; Glyco_hydro_18; 1.
DR   SMART; SM00494; ChtBD2; 1.
DR   PROSITE; PS01095; CHITINASE_18; 1.
KW   Glycosidase; Hydrolase.
SQ   SEQUENCE   472 AA;  52211 MW;  4234A6BEEFE3C165 CRC64;


  Query Match             86.4%;  Score 38;  DB 5;  Length 472;
  Best Local Similarity   77.8%;  Pred. No. 7.7;
  Matches    7;  Conservative    2;  Mismatches    0;  Indels    1;  Gaps

Qy       1 FGRSFTLAS 9
           :|:||||||
Db     279 YGRNFTLAS 287


RESULT   11
P91731
ID   P91731      PRELIMINARY;      PRT;   454 AA.
AC   P91731;
DT   01-MAY-1997 (TrEMBLrel. 03, Created)
DT   01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
```

```
ID   CHIT_CHAGR (truncated)  Reviewed;  [illegible] AA.
AC   [illegible];
DT   01-OCT-2001 (Rel. 40, Created)
DT   01-OCT-2001 (Rel. 40, Last sequence update)
DT   15-JUL-2003 (Rel. 44, Last annotation update)
DE   Chitinase precursor (EC 3.2.1.14) (Fragment).
GN   Name=CHIT;
OS   Chaoborus crystallinus (Phantom midge).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicomorpha;
OC   Chaoboridae; Chaoborus.
OX   NCBI_TaxID=7215;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Larva;
RX   MEDLINE=20120066; PubMed=11054542;
RA   Kramer K.J., Muthukrishnan S., Tomalski M.D., Kanost M.R., Arakane Y.,
RA   Zhu Y.C., Beeman R.W.;
RT   "Molecular cloning and sequence comparisons of chitinase genes among the
RT   insects and other species.";
RL   Insect Biochem. Mol. Biol. 31:129-141(2001).
CC   -!- SIMILARITY: BELONGS TO CHITINASE CLASS II (FAMILY 18 GLYCOSYL
CC       HYDROLASES).
DR   EMBL; U86677; AAB47504.1; -.
DR   InterPro; IPR001579; Chitinase_2.
DR   InterPro; IPR000726; Chitin_binding.
DR   InterPro; IPR001223; Glyco_hydro_18.
DR   Pfam; PF03067; Chitin_bind_2; 1.
DR   Pfam; PF00704; Glyco_hydro_18; 1.
DR   SMART; SM00494; ChtBD2; 1.
DR   PROSITE; PS01095; CHITINASE_18; 1.
KW   Glycosidase; Hydrolase.
SQ   SEQUENCE   554 AA;  [illegible] MW;  [illegible];
```

Query Match 86.4%; Score 48; DB 5; Length 554;
Best Local Similarity 77.8%; Pred. No. 8.1;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1 FGRSFTLAS 9
           :|| :|||:
Db     279 YGRSFTLAA 287
```

```
RESULT 12
Q9VZV2
ID   Q9VZV2      PRELIMINARY;      PRT:   696 AA.
AC   Q9VZV2;
DT   01-MAY-2001 (TrEMBLrel. 14, Created)
DT   01-MAY-2001 (TrEMBLrel. 14, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 17, Last annotation update)
DE   CG11869 PROTEIN.
GN   CG11869.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN BERKELEY
RX   MEDLINE=20190066; PubMed=10731132;
RA   Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA   Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA   George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA   Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA   Brandon R.C., Rogers Y.H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA   Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA   Abril J.F., Agbayani A., An H.J., Andrews-Pfannkoch C., Baldwin D.,
RA   Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA   Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA   Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA   Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA   Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA   de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA   Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C.,
RA   Dunn P., Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA   Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA   Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA   Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA   Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C.,
RA   Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA   Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA   Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA   Liu X., Mattei B., McIntosh T.C., Mcleod M.P., McPherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA   Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA   Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT   "The genome sequence of Drosophila melanogaster.";
RL   Science 287:2185-2195(2000).
DR   EMBL; AE003477; AAF47714.1; -.
DR   FlyBase; FBgn0035598; CG11869.
DR   InterPro; IPR001579; Chitinase_2.
DR   InterPro; IPR001223; Glyco_hydro_18.
DR   Pfam; PF00704; Glyco_hydro_18; 2.
DR   PROSITE; PS01095; CHITINASE_18; UNKNOWN_1.
SQ   SEQUENCE   696 AA;  77366 MW;  6877585837A19CP6 CRC64;
```

Query Match 86.4%; Score 48; DB 5; Length 696;
Best Local Similarity 77.8%; Pred. No. 10;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1 FGRSFTLAS 9
           :||| |||:
Db     370 YGRSFTLAN 378
```

```
RESULT 13
Q960M0
ID   Q960M0      PRELIMINARY;      PRT:  1013 AA.
AC   Q960M0;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   LD45559P.
GN   CG11869.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN Y, CN BW SP;
RA   Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA   Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,
RA   Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,
RA   Nunoo J., Pacleb J., Paragas V., Park S., Phouanenavong S., Wan K.,
RA   Yu C., Lewis S.E., Rubin G.M., Celniker S.;
RL   Submitted (AUG-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY051088; AAK93412.1; -.
SQ   SEQUENCE   1013 AA;  113265 MW;  26698IAD0685E0C2 CRC64;
```

Query Match 86.4%; Score 48; DB 5; Length 1013;
Best Local Similarity 77.8%; Pred. No. 16;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1 FGRSFTLAS 9
           :|||||||:
Db     370 YGRSFTLAN 378
```

```
RESULT  14
Q9V1Z4
ID   Q9V1Z4       PRELIMINARY;      PRT:   276 AA.
AC   Q9V1Z4;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   LETHAL (2) 37CC PROTEIN.
GN   L(2)37CC OR CG10691.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BERKELEY;
RX   MEDLINE=20109636; PubMed=10731132;
RA   Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA   Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA   George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA   Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA   Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA   Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA   Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA   Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA   Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA   Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA   Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA   Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA   de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA   Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA   Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA   Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA   Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA   Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA   Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA   Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA   Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA   Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA   Liu X., Mattei B., McIntosh T.C., Mcleod M.P., McPherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA   Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA   Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT   "The genome sequence of Drosophila melanogaster.";
RL   Science 287:2185-2195(2000).
DR   EMBL; AE003661; AAF53765.1; -.
DR   FlyBase; FBgn0002041; l(2)37Cc.
DR   InterPro; IPR001077; Band_7.
DR   InterPro; IPR000163; Prohibitin.
DR   Pfam; PF01145; Band_7; 1.
DR   PRINTS; PR00679; PROHIBITIN.
DR   SMART; SM00244; PHB; 1.
SQ   SEQUENCE   276 AA;  30384 MW;  EFE1CDB98A79F94F CRC64;


  Query Match             81.8%;  Score 36;  DB 5;  Length 276;
  Best Local Similarity   87.5%;  Pred. No. 10;
  Matches    7;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 FGRSFTLA 8
            | | ||||
Db     175 FGREFTLA 182


RESULT  15
Q96F97
ID   Q96F97       PRELIMINARY;      PRT:   390 AA.
AC   Q96F97;
DT   01-DEC-2001 (TrEMBLrel. 16, Created)
DT   01-DEC-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 16, Last annotation update)
DE   CHITINASE 4-LIKE 2.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=BRAIN, AND SMOOTH MUSCLE AORTA; MEDLINE=21301464;
RA   Strausberg R.;
RL   Submitted (OCT-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC011460; AAH11460.1; -.
SQ   SEQUENCE   390 AA;  43529 MW;  6B01E4AF7A1B63EC CRC64;


  Query Match             81.8%;  Score 36;  DB 4;  Length 390;
  Best Local Similarity   87.5%;  Pred. No. 13;
  Matches    7;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 FGRSFTLAS 9
            | ||||||
Db     269 YGHSFTLAS 277


Search completed: August 26, 2002, 06:28:34
Job time: 1030 sec
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

M protein protein search, using sw model

Run on: August 26, 2002, 06:38:58 ; Search time 1.28 Seconds
(without alignments)
7.177 Million cell updates/sec

Title: US-09-744-282-1
Perfect score: 44
Sequence: 1 FGPSFTLAS 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 74754 seqs, 11107496 residues

Total number of hits satisfying chosen parameters: 74754

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A Geneseq 03Aug02:*
1: /cgn2_6/ptodata/2/pubpep/geneseq/geneseq-pep-embl.dat:*
2: /cgn2_6/ptodata/2/pubpep/geneseq/geneseq-pep-embl.dat:*
3: /cgn2_6/ptodata/2/pubpep/geneseq/geneseq-pep-embl.dat:*
4: /cgn2_6/ptodata/2/pubpep/geneseq/geneseq-pep-embl.dat:*
5: /cgn2_6/ptodata/2/pubpep/geneseq/geneseq-pep-embl.dat:*
6: /cgn2_6/ptodata/2/pubpep/geneseq/geneseq-pep-embl.dat:*
7: /cgn2_6/ptodata/2/pubpep/geneseq/geneseq-pep-embl.dat:*
8: /cgn2_6/ptodata/2/pubpep/geneseq/geneseq-pep-embl.dat:*
9: /cgn2_6/ptodata/2/pubpep/geneseq/geneseq-pep-embl.dat:*
10: /cgn2_6/ptodata/2/pubpep/geneseq/geneseq-pep-embl.dat:*
11: /cgn2_6/ptodata/2/pubpep/geneseq/geneseq-pep-embl.dat:*
12: /cgn2_6/ptodata/2/pubpep/geneseq/geneseq-pep-embl.dat:*
13: /cgn2_6/ptodata/2/pubpep/geneseq/geneseq-pep-embl.dat:*
14: /cgn2_6/ptodata/2/pubpep/geneseq/geneseq-pep-embl.dat:*
15: /cgn2_6/ptodata/2/pubpep/geneseq/geneseq-pep-embl.dat:*
16: /cgn2_6/ptodata/2/pubpep/geneseq/geneseq-pep-embl.dat:*
17: /cgn2_6/ptodata/2/pubpep/geneseq/geneseq-pep-embl.dat:*
18: /cgn2_6/ptodata/2/pubpep/geneseq/geneseq-pep-embl.dat:*
19: /cgn2_6/ptodata/2/pubpep/geneseq/geneseq-pep-embl.dat:*
20: /cgn2_6/ptodata/2/pubpep/geneseq/geneseq-pep-embl.dat:*
21: /cgn2_6/ptodata/2/pubpep/geneseq/geneseq-pep-embl.dat:*
22: /cgn2_6/ptodata/2/pubpep/geneseq/geneseq-pep-embl.dat:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	44	100.0	9	17	AAB95071	Human cartilage gl
2	44	100.0	9	21	AAY69226	Peptide fragment o
3	44	100.0	9	17	AAB95075	Peptide conta. hC
4	44	100.0	9	21	AAB04489	Gp38k-derived cell
5	44	100.0	9	21	AAY69283	Peptide fragment o
6	44	100.0	9	19	AAW61598	Human Cartilage gp
7	44	100.0	9	18	AAW45139	Human cartilage gl
8	44	100.0	10	22	AAU02074	Synthetic human ta
9	44	100.0	22	22	AAU02078	Synthetic human mu
10	44	100.0	281	22	AAU02079	Synthetic multi ta
11	44	100.0	85	22	AAU02076	Synthetic human mu
12	44	100.0	388	21	AAB03442	Gp38k pr Gelin seq
13	44	100.0	459	22	AAU02077	Synthetic multi ta
14	41	93.2	373	19	AAW40262	Human chitinase pr
15	41	93.2	373	19	AAW40261	Human chitinase pr
16	41	93.2	373	20	AAY42427	Clone of the 39 ki
17	41	93.2	373	20	AAY42428	Chitinase amino ac
18	41	93.2	385	18	AAW26751	Bovine whey protei
19	41	93.2	387	18	AAW08585	Human 39 kDa chiti
20	41	93.2	464	18	AAW31497	Human chitotriosid
21	41	93.2	466	18	AAW31498	Human chitotriosid
22	41	93.2	466	18	AAW08584	Human 50 kDa chiti
23	41	93.2	466	19	AAW40259	Human chitinase pr
24	41	93.2	466	19	AAW40260	Human chitinase pr
25	41	93.2	466	20	AAY42425	Mr 21k clone of hu
26	41	93.2	466	20	AAY42426	Mr 24k clone of hu
27	41	93.2	466	22	AAB00442	Human chitinase pr
28	41	93.2	466	22	AAB00443	Human chitinase pr
29	38	86.4	16	18	AAW45140	Human cartilage gl
30	38	86.4	16	19	AAW61599	Human Cartilage gp
31	38	86.4	199	22	ABB56452	Human sitaxic chain
32	38	86.4	199	22	ABB56456	Human sitaxic chain
33	38	86.4	208	22	ABB56475	Human sitaxic chain
34	38	86.4	210	22	ABB56442	Human sitaxic chain
35	38	86.4	211	22	ABB56441	Human sitaxic chain
36	38	86.4	211	22	ABB56455	Human sitaxic chain
37	38	86.4	212	22	ABB56439	Human sitaxic chain
38	38	86.4	212	22	ABB56440	Human sitaxic chain
39	38	86.4	212	22	ABB56477	Human sitaxic chain
40	38	86.4	212	22	AAU51698	Propionibacterium
41	38	86.4	213	22	ABB56443	Human sitaxic chain
42	38	86.4	213	22	ABB56445	Human sitaxic chain
43	38	86.4	213	22	ABB56446	Human sitaxic chain
44	38	86.4	213	22	ABB56453	Human sitaxic chain
45	38	86.4	214	22	ABB56444	Human sitaxic chain

ALIGNMENTS

RESULT 1
AAB95071
ID AAB95071 standard; peptide; 9 AA.
XX
AC AAB95071;
XX
DT [illegible]
XX
DE Human cartilage glycoprotein 39 autoantigen derived peptide (.).
XX
KW human; cartilage; glycoprotein 39; hC gp-39; autoantigen;
KW induction, T cell, tolerance, mediated, destruction;
KW rheumatoid arthritis, disease, model, diagnosis.
XX
OS Homo sapiens.
XX
PN WO9613517-A1.
XX
PD 09-MAY-1996.
XX
PF 25-OCT-1995; 95WO-EP04201.
XX
PR 07-APR-1995; 95EP-0200886.
PR 27-OCT-1994; 94EP-0203128.
XX
PA (AKZO) AKZO NOBEL NV.
XX
PI Boots AMH, Verheijden GFM;
XX
DR WPI; 1996-239447/24.
XX
PT Peptide(s) derived from human cartilage gp-39 autoantigen, - used
PT to induce specific T cell tolerance, esp. for treatment of
PT rheumatoid arthritis, also usetul in producing arthritic animal

```
PT   models
XX
PS   Claim 1; Page 28; 47pp; English.
XX
CC   The present peptide, derived from the human cartilage glycoprotein
CC   39 (hC gp 39) autoantigen, can be used to induce specific T-cell
CC   tolerance to hC gp 39 in a patient with T-cell mediated cartilage
CC   destruction, esp. rheumatoid arthritis (RA). It may also be used to
CC   induce RA in mice, etc. (human disease models), and to diagnose the
CC   presence of activated T cells autoreactive with hC gp 39, i.e. to
CC   identify patients suitable for tolerising treatment.
CC   Therapeutically the peptide is given in a dose of 0.1 100 microg/kg
CC   by injection, or 10 50 microg/kg to produce a disease model. The
CC   peptide has a specific tolerising effect on autoreactive T cells,
CC   but leaves other components of the immune system intact, and
CC   therefore should have few side effects. In an assay to determine
CC   the peptide's binding to HLA-DR mols., specifically DR4Dw4 (known
CC   to be associated with an increased risk of RA development), the
CC   peptide gave a D-I Plot IC(50) (micromolar) value of 0.04 compared
CC   to 0.55 for the control peptide IRA(307-319)F.
XX
SQ   Sequence   9 AA;


  Query Match             100.0%;  Score 44;  DB 17;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 6.4e-05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 FGRSFTLAS 9
           |||||||||
Db       1 fgrsftlas 9


RESULT   2
AAY69226
ID   AAY69226 standard; peptide; 9 AA.
XX
AC   AAY69226;
XX
DT   30-MAY-2000  (first entry)
XX
DE   Peptide treatment of human cartilage gp-39 (HC gp 39) protein.
XX
KW   Human cartilage gp 39 protein; HC gp 39 protein; T cell activator;
KW   inflammatory disease; lymphocyte; antigen; bystander suppression;
KW   antigen specific T cell tolerance; autoimmune disorder;
KW   inflammatory disease; rheumatoid arthritis;
KW   autoimmune inflammatory disease; autoantigen.
XX
OS   Homo sapiens.
XX
PN   WO200004917-A2.
XX
PD   03-FEB-2000.
XX
PF   19-JUL-1999;   99WO-EP05331.
XX
PR   23-JUL-1998;   98EP-0202471.
XX
PA   (ALKU ) AKZO NOBEL NV.
XX
PI   Millenburg AMM;  Boots AMB;
XX
DR   WPI; 2000-182547/16.
XX
PT   Use of cartilage protein or peptides for treating inflammatory diseases
PT   especially rheumatoid arthritis
XX
PS   Claim 2; Page 17; 29pp; English.
XX
CC   AAY69226-AB represent treatments of the human cartilage gp-39 (HC gp 39)
CC   protein. The peptides are T cell activators. The peptides are used for
CC   preventing inflammatory diseases, by modulating lymphocyte reactive
CC   to antigens other than HC gp 39, present in the same tissue as
CC   HC gp 39, by the induction of antigen specific T cell tolerance.
CC   autoimmune disorders can be treated by bystander suppression. The
CC   HC gp 39 peptides are useful for preventing inflammatory diseases
CC   especially rheumatoid arthritis. The treatments may also be used
CC   treating autoimmune inflammatory diseases. The HC gp 39 peptides are
CC   especially beneficial in autoimmune conditions in which the
CC   autoantigens associated with the condition are unknown.
XX
SQ   Sequence   9 AA;


  Query Match             100.0%;  Score 44;  DB 21;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 6.4e-05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 FGRSFTLAS 9
           |||||||||
Db       1 fgrsftlas 9


RESULT   3
AAR95075
ID   AAR95075 standard; peptide; 13 AA.
XX
AC   AAR95075;
XX
DT   17-DEC-1996  (first entry)
XX
DE   Peptide contg. hC gp 39 autoantigen derived peptide (1).
XX
KW   Human; cartilage; glycoprotein; hC gp 39; autoantigen;
KW   induction; T cell; tolerance; mediated; destruction;
KW   rheumatoid arthritis; disease model; diagnosis.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Peptide          3..11
FT                    /note= "HC gp 39 peptide (1)"
XX
PN   WO9614517-A1.
XX
PD   09-MAY-1996.
XX
PF   25-OCT-1995;   95WO-EP04241.
XX
PR   07-APR-1995;   95EP-0200886.
PR   27-OCT-1994;   94EP-0203128.
XX
PA   (ALKU ) AKZO NOBEL NV.
XX
PI   Boots AMB;  Verheijden GPM;
XX
DR   WPI; 1996-239447/24.
XX
PT   Peptide(s) derived from human cartilage gp 39 autoantigen -
PT   to induce specific T cell tolerance, esp. for treatment of
PT   rheumatoid arthritis, also useful in predicting arthritis animal
PT   models
XX
PS   Claim 2; Page 28; 47pp; English.
XX
CC   The present peptide, contg. a peptide derived from the human
CC   cartilage glycoprotein 39 (hC gp 39) autoantigen, can be used t
CC   induce specific T cell tolerance to hC gp 39 in a patient with
CC   T-cell mediated cartilage destruction, esp. rheumatoid arthrit
CC   (RA). It may also be used to induce RA in mice, etc. (human dis
CC   models), and to diagnose the presence of activated T cells
CC   autoreactive with hC gp 39, i.e. to identify patients suitable f
CC   tolerising treatment.
CC   Therapeutically the peptide is given in a dose of 0.1 100 micro
CC   by injection, or 10 50 microg/kg to produce a disease model. The
```

```
CC   peptide has a specific toxic effect on autoreactive T-cells,
CC   but leaves other components of the immune system intact, and
CC   therefore should have few side effects.
XX
SQ   Sequence   15 AA;


  Query Match             100.0%;  Score 44;  DB 17;  Length 15;
  Best Local Similarity   100.0%;  Pred. No. 0.016;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 FGRSFTLAS 9
           ||| ||||| |
Db       3 fgrsftlas 11


RESULT   4
AAB03439
ID   AAB03439 standard; Peptide; 15 AA.
XX
AC   AAB03439;
XX
DT   28-JAN-2001  (first entry)
XX
DE   Gp38k derived cell migration peptide #5.
XX
KW   Gp38k; chemoattractant; cell migration; wound healing; angiogenesis;
KW   cancer; vascular trauma; vascular disease; atherosclerosis; restenosis.
XX
OS   Unidentified.
XX
PN   WO200034469-A1.
XX
PD   15-JUN-2000.
XX
PF   10-DEC-1999;   99WO-US29262.
XX
PR   11-DEC-1998;   98US-0111850.
XX
PA   (UYNY ) UNIV NEW YORK STATE RES FOUND.
XX
PI   Mullis AJ;
XX
DR   WPI; 2000-438130/37.
XX
PT   Clusterin and gp38K derived peptide capable of altering cell migration
PT   useful for treating atherosclerosis, cancer and stenosis following
PT   vascular trauma or disease
XX
PS   Claim 1; Page 27; 49pp; English.
XX
CC   The present sequence is a peptide derived from gp38k, Gp38k is
CC   essential for the migration of vascular smooth muscle cells (VSMC). The
CC   gene and protein can therefore be used to promote wound healing,
CC   anti-stenosis and vasculogenesis, in the treatment of stenosis following
CC   vascular trauma or disease and to treat atherosclerosis, and antisense
CC   sequences can be used to treat cancer, as angiogenesis is vital for
CC   tumour survival. This peptide can be used to alter cell migration
CC   patterns.
XX
SQ   Sequence   15 AA;


  Query Match             100.0%;  Score 44;  DB 21;  Length 15;
  Best Local Similarity   100.0%;  Pred. No. 0.016;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 FGRSFTLAS 9
           |||||||||
Db       1 fgrsftlas 9


RESULT   5
AAY69230
ID   AAY69230 standard; peptide; 13 AA.
XX
AC   AAY69230;
XX
DT   30-MAY-2000  (first entry)
XX
DE   Peptide fragment of human cartilage gp-39 (HC gp-39) protein.
XX
KW   Human cartilage gp-39 protein; HC gp-39 protein; T-cell activator;
KW   inflammatory disease; lymphocyte; antigen; bystander suppression;
KW   antigen-specific T-cell tolerance; autoimmune disorder;
KW   inflammatory disease; rheumatoid arthritis;
KW   autoimmune inflammatory disease; autoantigen.
XX
OS   Homo sapiens.
XX
PN   WO200004917-A2.
XX
PD   03-FEB-2000.
XX
PF   19-JUL-1999;   99WO-EP05331.
XX
PR   23-JUL-1998;   98EP-0202471.
XX
PA   (AKZO ) AKZO NOBEL NV.
XX
PI   Miltenburg AMM, Boots AMH;
XX
DR   WPI; 2000-182547/16.
XX
PT   Use of cartilage protein or peptides for treating inflammatory diseases
PT   especially rheumatoid arthritis -
XX
PS   Claim 2; Page 17; 29pp; English.
XX
CC   AAY69226-33 represent fragments of the human cartilage gp-39 (HC gp-39)
CC   protein. The peptides are T-cell activators. The peptides are used for
CC   preventing inflammatory diseases, by modulating lymphocytes reactive
CC   to antigens other than HC gp-39, present in the same tissue of
CC   HC gp-39. By the induction of antigen-specific T-cell tolerance,
CC   autoimmune disorders can be treated by bystander suppression. The
CC   HC gp-39 peptides are useful for preventing inflammatory diseases
CC   especially rheumatoid arthritis. The fragments may also be used for
CC   treating autoimmune inflammatory diseases. The HC gp-39 peptides are
CC   especially beneficial in autoimmune conditions in which the
CC   autoantigens associated with the condition are unknown.
XX
SQ   Sequence   13 AA;


  Query Match             100.0%;  Score 44;  DB 21;  Length 13;
  Best Local Similarity   100.0%;  Pred. No. 0.016;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 FGRSFTLAS 9
           |||||||||
Db       3 fgrsftlas 11


RESULT   6
AAW61598
ID   AAW61598 standard; protein; 15 AA.
XX
AC   AAW61598;
XX
DT   27-OCT-1998  (first entry)
XX
DE   Human Cartilage gp-39 epitope (ii).
XX
KW   Human; cartilage gp-39; monoclonal antibody; cell surface antigen;
KW   rheumatoid arthritis; T-cell receptor.
XX
```

OS Homo sapiens.
XX
PN EP0856520-A1.
XX
PD 05-AUG-1998.
XX
PF 02-DEC-1997; 97EP-0203769.
XX
PR 27-JUN-1997; 97EP-0201972.
PR 06-DEC-1996; 96EP-0203465.
XX
PA (ALEX) AKZO NOBEL NV.
PA (ALKU) AKZO NV.
XX
PI Stroobakkers HGA;
XX
DR WPI; 1998-400866/35.
XX
PT Preparation of monoclonal antibodies against cell surface antigens
PT used as diagnostic agents and to treat rheumatoid arthritis -
PT comprises immunising, then enriching, enriching and cloning B cells
PT and then selecting and cloning hybridoma
XX
PS Examples; Page 14; 24pp; English.
XX
CC The human cartilage gp-39 peptides AAW61597-W61599 were used in the
CC preparation of a monoclonal antibody against a cell surface antigen.
CC The antibody can be used for treatment of rheumatoid arthritis. It
CC is also useful against cell surface antigens and antibodies reactive
CC with the clonotypic structure of T cell receptors are used as diagnostic
CC reagents. The process allows antibodies to be raised against
CC antigens which are present in very low amounts or which are
CC conformationally unstable without the need to screen very large numbers
CC of hybridomas and without requiring high purity Antigen for detachment.
XX
SQ Sequence 15 AA;

Query Match 100.0%; Score 44; DB 19; Length 15;
Best Local Similarity 100.0%; Pred. No. 0.019;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FGKSFTLAS 9
||||||||| |
Db 1 fgksftlas 9

RESULT 7
AAW45149
ID AAW45149 standard; peptide; 29 AA.
XX
AC AAW45149;
XX
DT 28-APR-1998 (first entry)
XX
DE Human cartilage glycoprotein 39 derived peptide #40.
XX
KW Articular cartilage; immunosuppressive therapy; antisera; antiantigen;
KW immunological tolerance; T cell; human cartilage glycoprotein 39;
KW HC gp 39; rheumatoid arthritis; epitope.
XX
OS Synthetic.
OS Homo sapiens.
XX
PN WO9740068-A1.
XX
PD 30-OCT-1997.
XX
PF 22-APR-1997; 97WO-EP02051.
XX
PR 24-APR-1996; 96EP-0201106.
XX
PA (ALKO) AKZO NOBEL NV.
XX
PI Boots AMH, Verboijden GFM;
XX
DR WPI; 1997-545773/49.
XX
PT Peptide suitable for use in antigen specific immunosuppressive
PT therapy - resembles or mimics epitope present on HC gp-39, is
PT inducing systemic immunological tolerance to rheumatoid arthritis
PT autoantigen
XX
PS Disclosure; Page 16; 82pp; English.
XX
CC The present sequence represents a peptide which resembles or mimics an
CC epitope present in human cartilage glycoprotein 39 (HC gp-39), is
CC autoantigen in rheumatoid arthritis. The invention relates to peptides
CC consisting of 10-55 amino acid residues comprising at least one of the
CC following 19 sequences: LVGYYISWS; FDPTHTIYS; TIYSFANIS; IKITLVNS;
CC FIKSVPPFL; FPSLFAWL; LYFPEGKS; YDCAKINQH; LDFISIMIY; EISIMIYSI;
CC FRQGDASS; YAVGYMLFL; MSSCVASS; LAYYEIDI; IEGAIVHH; YIKGGLIA;
CC LAGAMVWAL; VWALDLDDF; or LDLDDFQGS. They can be used peptides in
CC antigen specific immunosuppressive therapy, particularly the prevention
CC of T-cell mediated destruction of articular cartilage in autoimmune
CC diseases (e.g. rheumatoid arthritis). They can also be used to detect
CC activated autoreactive T cells in an individual. The peptides have a
CC specific effect on the autoreactive T cells, thus leaving the other
CC components of the immune system intact, unlike the non-specific
CC suppressive effect of immunosuppressive drugs, and do not cause side
CC side effects. The peptides are predominantly recognised by activated
CC T cells from rheumatoid arthritis patients, but rarely by those from
CC healthy donors.
XX
SQ Sequence 16 AA;

Query Match 100.0%; Score 44; DB 18; Length 16;
Best Local Similarity 100.0%; Pred. No. 0.022;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FGKSFTLAS 9
||||||||| |
Db 6 fgksftlas 14

RESULT 8
AAU02074
ID AAU02074 standard; Protein; 305 AA.
XX
AC AAU02074;
XX
DT 07-SEP-2001 (first entry)
XX
DE Synthetic human target autoantigen sheLEP/RA.
XX
KW sheLEP/RA; immunogenic epitope cluster; IEC;
KW synthetic human chondrocyte glycoprotein 39/cartilage link protein;
KW autoantigen; autoimmune disease; multiple sclerosis;
KW rheumatoid arthritis; insulin-dependent diabetes mellitus;
KW myasthenia gravis; uveitis; autoimmune hepatitis; thyroiditis; orchitis;
KW orchitis; idiopathic thrombocytopenic purpura; inflammatory disease;
KW Crohn's disease; ulcerative colitis.
XX
OS Synthetic.
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Misc-difference 1..5
FT /label= CHEF
FT /note= "Other STOP"
XX
PN WO200141037-A2.
XX
PD 07-JUN-2001.
XX

```
PF   26-OCT-2000; 2000WO-IL00688.
XX
PR   27-OCT-1999;  99IL-132611.
XX
PA   (YEDA ) YEDA RES & DEV CO LTD.
XX
PI   Ben-Nun A,  Kerlero de Rosbo N,  Sappler GP;
XX
DR   WPI; 2001-300515/31.
DR   N-PSDB; AAS04859.
XX
PI   Novel synthetic human target autoantigen gene useful for treating
PI   autoimmune diseases such as multiple sclerosis, insulin dependent
PI   diabetes mellitus, rheumatoid arthritis, myasthenia gravis, and uveitis
PI
XX
PS   Claim 185; Fig 59; 182pp; English.
XX
CC   The sequence represents synthetic human multi-target autoantigen
CC   (shMultiTAG) Y-EAPd consisting of regions encoding preferred epitopes
CC   from shCollagen%A -hAgrin a-%A and shMT-%A. The synthetic human
CC   target autoantigen genes of the invention comprise sequences coding for
CC   at least 2 IECs of autoantigen(s) related to a specific autoimmune
CC   disease. The synthetic human target autoantigen genes are useful for
CC   treating autoimmune diseases such as multiple sclerosis, insulin
CC   dependent diabetes mellitus, rheumatoid arthritis, myasthenia gravis,
CC   uveitis, autoimmune hepatitis, thyroiditis, insulitis, orchitis,
CC   idiopathic thrombocytopenic purpura, and inflammatory diseases (Crohn's
CC   disease, ulcerative colitis). The synthetic human target autoantigen
CC   genes are also useful for diagnosis and/or monitoring the progression of
CC   the autoimmune disease.
XX
SQ   Sequence  220 AA;


  Query Match             100.0%;  Score 44;  DB 22;  Length 220;
  Best Local Similarity   100.0%;  Pred. No. 0.31;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 FGRSFILAS 9
           |||||||||
Db     183 fgrsfilas 191


RESULT  10
AA002079
ID   AA002079 standard; protein; 281 AA.
XX
AC   AA002079;
XX
DT   07-SEP-2001  (first entry)
XX
DE   Synthetic multi-target autoantigen Y-EAPd.
XX
KW   Y-EAPd; immunogenic epitope cluster; IEC;
KW   autoimmune disease; multiple sclerosis; rheumatoid arthritis;
KW   insulin-dependent diabetes mellitus; myasthenia gravis; uveitis;
KW   autoimmune hepatitis; thyroiditis; insulitis; orchitis;
KW   idiopathic thrombocytopenic purpura; inflammatory disease;
KW   Crohn's disease; ulcerative colitis; shMultiTAG.
XX
OS   Synthetic.
OS   Homo sapiens.
OS   Escherichia coli.
OS   Mycobacterium tuberculosis.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 279
FT                   /label= OTHER
FT                   /note= "Other- STOP"
XX
```

```
PN   WO200131037-A2.
XX
PD   03 MAY 2001.
XX
PF   26 OCT 2000; 2000WO-IL00688.
XX
PR   27 OCT 1999;  99IL-013271.
XX
PA   (YEDA ) YEDA RES & DEV CO LTD.
XX
PI   Ben-Nun A,  Kerlero De Rosbo N,  Sappler GP;
XX
DR   WPI; 2001-300515/31.
DR   N-PSDB; AAS04837.
XX
PT   Novel synthetic human target autoantigen gene useful for treating
PT   autoimmune diseases such as multiple sclerosis, insulin-dependent
PT   diabetes mellitus, rheumatoid arthritis, myasthenia gravis, and uveitis
PT
XX
PS   Claim 179; Fig 55; 182pp; English.
XX
CC   The sequence represents synthetic human multi target autoantigen
CC   (shMultiAG) Y-RAPd consisting of preferred epitopes encoded by
CC   shSclAgen/KA, shAquaporin/KA and shCII/KA. The synthetic human
CC   target autoantigen genes of the invention comprise sequences coding for
CC   at least 2 IE's of autoantigen(s) related to a specific autoimmune
CC   disease. The synthetic human target autoantigen genes are useful for
CC   treating autoimmune diseases such as multiple sclerosis, insulin-
CC   dependent diabetes mellitus, rheumatoid arthritis, myasthenia gravis,
CC   uveitis, autoimmune hepatitis, thyroiditis, insulitis, orchitis,
CC   idiopathic thrombocytopenic purpura, and inflammatory diseases (Crohn's
CC   disease, ulcerative colitis). The synthetic human target autoantigen
CC   genes are also useful for diagnosis and/or monitoring the progression of
CC   the autoimmune disease.
XX
SQ   Sequence   281 AA;
```

```
  Query Match             100.0%;  Score 44;  DB 22;  Length 281;
  Best Local Similarity   100.0%;  Pred. No. 0.4;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 PGRSFTLAS 9
           |||||||||
Db     183 PGRSFTLAS 191
```

```
RESULT  11
AAU02076
ID   AAU02076 standard; Protein; 350 AA.
XX
AC   AAU02076;
XX
DT   07-SEP-2001  (first entry)
XX
DE   Synthetic human multi target autoantigen Y-RAPa.
XX
KW   Y-RAPa; immunogenic epitope cluster; IEC;
KW   autoantigen; shSclAgen/KA; shAquaporin/KA; shCII/KA;
KW   autoimmune disease; multiple sclerosis; rheumatoid arthritis;
KW   insulin dependent diabetes mellitus; myasthenia gravis; uveitis;
KW   autoimmune hepatitis; thyroiditis; insulitis; orchitis;
KW   idiopathic thrombocytopenic purpura; inflammatory disease;
KW   Crohn's disease; ulcerative colitis; shMultiAG.
XX
OS   Synthetic.
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Misc difference 348
FT                   /label= OTHER
FT                   /note= "other  STOP"
```

```
XX
PN   WO200131037-A2.
XX
PD   03 MAY 2001.
XX
PF   26-OCT-2000; 2000WO-IL00688.
XX
PR   27-OCT-1999;  99IL-013271.
XX
PA   (YEDA ) YEDA RES & DEV CO LTD.
XX
PI   Ben-Nun A,  Kerlero De Rosbo N,  Sappler GP;
XX
DR   WPI; 2001-300515/31.
DR   N-PSDB; AAS04837.
XX
PT   Novel synthetic human target autoantigen gene useful for treating
PT   autoimmune diseases such as multiple sclerosis, insulin dependent
PT   diabetes mellitus, rheumatoid arthritis, myasthenia gravis, and uveitis
PT
XX
PS   Claim 179; Fig 55; 182pp; English.
XX
CC   The sequence represents synthetic human multi target autoantigen
CC   (shMultiAG) Y-RAPa consisting of peptides encoded by shSclAgen/KA,
CC   shAquaporin/KA and shCII/KA. The synthetic human target autoantigen genes
CC   of the invention comprise sequences coding for at least 2 IE's of
CC   autoantigen(s) related to a specific autoimmune disease. The synthetic
CC   human target autoantigen genes are useful for treating autoimmune
CC   diseases such as multiple sclerosis, insulin dependent diabetes
CC   mellitus, rheumatoid arthritis, myasthenia gravis, uveitis, autoimmune
CC   hepatitis, thyroiditis, insulitis, orchitis, idiopathic thrombocytopenic
CC   purpura, and inflammatory diseases (Crohn's disease, ulcerative colitis).
CC   The synthetic human target autoantigen genes are also useful for
CC   diagnosis and/or monitoring the progression of the autoimmune disease.
XX
SQ   Sequence   350 AA;
```

```
  Query Match             100.0%;  Score 44;  DB 22;  Length 350;
  Best Local Similarity   100.0%;  Pred. No. 0.5;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 PGRSFTLAS 9
           |||||||||
Db     272 PGRSFTLAS 280
```

```
RESULT  12
AAB03442
ID   AAB03442 standard; Protein; 383 AA.
XX
AC   AAB03442;
XX
DT   03-JAN-2001  (first entry)
XX
DE   Gp38k protein sequence.
XX
KW   Gp38k; chemoattractant; cell migration; wound healing; angiogenesis;
KW   cancer; vascular trauma; vascular disease; atherosclerosis; restenosis;
XX
OS   Unidentified.
XX
FH   Key             Location/Qualifiers
FT   Peptide         1..24
FT                   /label= signal peptide
FT   Region          61..87
FT                   /note= "peptide antibody"
FT   Modified site   59..61
FT                   /label= glycosylation site
FT   Binding site    68..75
FT                   /label= leucine zipper
FT   Binding site    82..89
```

```
FT                     /label= zinc_finger_C2H2_type
FT   Active site       1..46
FT                     /label= chitinase_active_site
FT   Binding site      1..46
FT                     /label= heparin_binding_site
FT   Binding site      1..64
FT                     /label= hyaluronic_acid_binding_site
FT   Binding site      2..270
FT                     /label= hyaluronic_acid_binding_site
FT   Binding site      2..281
FT                     /label= oligosaccharide_binding_site
FT   Region            3..357
FT                     /note= "acidic region"
FT   Binding site      5..376
FT                     /label= hyaluronic_acid_binding_site
XX
PN   WO2001034644-A1.
XX
PD   17-MAY-2001.
XX
PF   10-NOV-1999;  99WO-US29262.
XX
PR   11-NOV-1998;  98US-0191856.
XX
PA   (UYNY ) UNIV NEW YORK STATE RES FOUND.
XX
PI   Mikis AM;
XX
DR   WPI; 2001-344000/37.
XX
PT   Clusterin and apoE-derived peptide capable of altering cell migration
PT   useful for treating atherosclerosis, cancer and stenosis following
PT   vascular trauma or disease
XX
PS   Disclosure; Fig 1; 44pp; English.
XX
CC   The present sequence is the protein sequence of apoEdp, ApoEdp, a
CC   chemoattractant, is essential for the migration of vascular smooth muscle
CC   cells (VSMC). The peptide is useful for treating wound healing,
CC   wound healing, angiogenesis and vasculogenesis, in the treatment of
CC   stenosis following vascular trauma or disease and to treat
CC   atherosclerosis, and antisense sequences can be used to treat cancer, as
CC   angiogenesis is vital to tumour survival.
XX
SQ   Sequence  383 AA;


  Query Match             100.0%;  Score 44;  DB 21;  Length 383;
  Best Local Similarity   100.0%;  Pred. No. 0.55;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 FGRSFTLAS 9
             |||||||||
Db     261 fgrsftlas 269


RESULT  13
AAB02277
ID   AAB02277 standard; Protein; 459 AA.
XX
AC   AAB02277;
XX
DT   27-SEP-2001  (first entry)
XX
DE   Synthetic multi-target autoantigen Y-hAPa.
XX
KW   Y-hAPa; immunogenic epitope cluster; IEC;
KW   autoantigen; shChitinase/3A; 3-type; rA/3A; rA/3E/3A; rA/3C/3A;
KW   autoimmune disease; multiple sclerosis; rheumatoid arthritis;
KW   insulin dependent diabetes mellitus; myasthenia gravis; uveitis;
KW   autoimmune hepatitis; thyroiditis; insulitis; orchitis;
KW   idiopathic thrombocytopenic purpura; inflammatory disease;
KW   Crohn's disease; ulcerative colitis; shMultiTAG.
XX
OS   Synthetic.
OS   Homo sapiens.
OS   Escherichia coli.
OS   Mycobacterium tuberculosis.
XX
FH   Key             Location/Qualifiers
FT   Misc difference 455
FT                   /label= OTHER
FT                   /note= "other: STOP"
FT   Misc-difference 457
FT                   /label= OTHER
FT                   /note= "other: STOP"
XX
PN   WO200134037-A2.
XX
PD   03-MAY-2001.
XX
PF   26-OCT-2000; 2000WO-IL00688.
XX
PR   27-OCT-1999;  99IL-0132611.
XX
PA   (YEDA ) YEDA RES & DEV CO LTD.
XX
PI   Ben Nun A, Kerlero De Rosbo N, Sappler CP;
XX
DR   WPI; 2001-330515/34.
DR   N-PSDB; AAS04838.
XX
PT   Novel synthetic human target autoantigen gene useful for treating
PT   autoimmune diseases such as multiple sclerosis, insulin dependent
PT   diabetes mellitus, rheumatoid arthritis, myasthenia gravis, and uveitis
PT   .
XX
PS   Claim 182; Fig 57; 182pp; English.
XX
CC   The sequence represents synthetic human multi target autoantigen
CC   (shMultiTAG) Y-hAPa consisting of regions encoding shChitinase/3A,
CC   shChitinase/3A, shMIF/3A and shGST/3A. The synthetic human target
CC   autoantigen genes of the invention comprise sequences coding for at least
CC   2 IECs of autoantigen(s) related to a specific autoimmune disease. The
CC   synthetic human target autoantigen genes are useful for treating
CC   autoimmune diseases such as multiple sclerosis, insulin dependent
CC   diabetes mellitus, rheumatoid arthritis, myasthenia gravis, uveitis,
CC   autoimmune hepatitis, thyroiditis, insulitis, orchitis, idiopathic
CC   thrombocytopenic purpura, and inflammatory diseases (Crohn's disease,
CC   ulcerative colitis). The synthetic human target autoantigen genes are
CC   also useful for diagnosis and/or monitoring the progression of the
CC   autoimmune disease.
XX
SQ   Sequence  459 AA;


  Query Match             100.0%;  Score 44;  DB 22;  Length 459;
  Best Local Similarity   100.0%;  Pred. No. 0.66;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 FGRSFTLAS 9
             |||||||||
Db     272 fgrsftlas 280


RESULT  14
AAW40262
ID   AAW40262 standard; Protein; 374 AA.
XX
AC   AAW40262;
XX
DT   15-JUN-1998  (first entry)
XX
DE   Human chitinase protein analogue.
XX
KW   Chitinase; human; fungal infection; immunogen; diagnosis; treatment;
```

```
KW   Gaucher's disease; transgenic; detection; hybridisation; antifungal;
KW   rheumatoid arthritis; overexpression; extracellular matrix.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Misc difference 370
FT                   /label= p370S
FT                   /note= "Wild type Pro is replaced by Ser"
XX
PN   WO9747752-A1.
XX
PD   18-DEC-1997.
XX
PF   16-JUN-1997;    97WO-US010460.
XX
PR   14-JUN-1996;    96US-0663618.
XX
PA   (ICOS ) ICOS CORP.
XX
PI   Gray PW;
XX
DR   WPI; 1998-052316/05.
XX
PT   Nucleic acids encoding human chitinase - useful as antifungal
PT   agents, especially in combination with other antifungals
XX
PS   Claim 29; Page 44-50; 63pp; English.
XX
CC   This sequence encodes a fragment of a novel human chitinase protein
CC   analogue in which a proline at position 370 of the wild type protein
CC   (see AAW40261) is substituted with a serine and the C terminal 72
CC   residues of the mature protein are deleted. These residues are not
CC   critical to chitinase enzymatic activity. Chitinases are useful for
CC   treating or preventing fungal infection and as immunogens for generating
CC   antibodies which are used to purify, detect and quantify chitinases, e.g.
CC   for diagnosis of Gaucher's disease. The nucleic acid sequence of the
CC   chitinase is also useful as a probe to identify and isolate genomic DNA
CC   encoding chitinases or similar proteins, or cells expressing them or to
CC   generate transgenic ('knockout') rodents. It can also be used in
CC   hybridisation assays and to detect genetic alterations in the chitinase
CC   gene related to disease. Agents that inhibit this protein may be useful
CC   in treatment of Gaucher's disease and rheumatoid arthritis, where
CC   overexpression of the protein can damage the extracellular matrix.
CC   Chitinase also improves the activity of other antifungal agents and may
CC   allow a reduction in the dose of such agents, and thus of their side
CC   effects.
XX
SQ   Sequence   473 AA;


  Query Match             93.2%;  Score 41;  DB 19;  Length 473;
  Best Local Similarity   88.9%;  Pred. No. 2.2;
  Matches    8;  Conservative    1;  Mismatches    0;  Indels    1;  Gaps    0;

Qy       1 FGRSFTLAS 9
             :|||:||||
Db     246 yarsttlas 254


RESULT  15
AAW40261
ID   AAW40261 standard; Protein; 473 AA.
XX
AC   AAW40261;
XX
DT   15-JUN-1998  (first entry)
XX
DE   Human chitinase protein fragment.
XX
KW   Chitinase; human; fungal infection; immunogen; diagnosis; treatment;
KW   Gaucher's disease; transgenic; detection; hybridisation; antifungal;
KW   rheumatoid arthritis; overexpression; extracellular matrix.
XX
OS   Homo sapiens.
XX
PN   WO9747752-A1.
XX
PD   18-DEC-1997.
XX
PF   16-JUN-1997;    97WO-US010460.
XX
PR   14-JUN-1996;    96US-0663618.
XX
PA   (ICOS ) ICOS CORP.
XX
PI   Gray PW;
XX
DR   WPI; 1998-052316/05.
DR   N-PSDB; AAV10435.
XX
PT   Nucleic acids encoding human chitinase - useful as antifungal
PT   agents, especially in combination with other antifungals
XX
PS   Claim 29; Page 48-49; 63pp; English.
XX
CC   This sequence encodes a fragment of a novel human chitinase protein
CC   which lacks the C terminal 72 residues of the mature protein. These
CC   residues are not critical to chitinase enzymatic activity. Chitinases are
CC   useful for treating or preventing fungal infection and as immunogens for
CC   generating antibodies which are used to purify, detect and quantify
CC   chitinases, e.g. for diagnosis of Gaucher's disease. The nucleic acid
CC   sequence of the chitinase is also useful as a probe to identify and
CC   isolate genomic DNA encoding chitinases or similar proteins, or cells
CC   expressing them or to generate transgenic ('knockout') rodents. It can
CC   also be used in hybridisation assays and to detect genetic alterations
CC   in the chitinase gene related to disease. Agents that inhibit this
CC   protein may be useful in treatment of Gaucher's disease and rheumatoid
CC   arthritis, where overexpression of the protein can damage the
CC   extracellular matrix. Chitinase also improves the activity of other
CC   antifungal agents and may allow a reduction in the dose of such agents
CC   and thus of their side effects.
XX
SQ   Sequence   473 AA;


  Query Match             93.2%;  Score 41;  DB 19;  Length 473;
  Best Local Similarity   88.9%;  Pred. No. 2.2;
  Matches    8;  Conservative    1;  Mismatches    0;  Indels    1;  Gaps    0;

Qy       1 FGRSFTLAS 9
             :|||:||||
Db     246 yarsttlas 254


Search completed: August 26, 2002, 06:13:24
Job time: 256 sec
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 26, 2002, 05:05:56 ; Search time 40.22 Seconds
(without alignments)
4.756 Million cell updates/sec

Title: US-09-744-282-1
Perfect score: 44
Sequence: 1 FGRSFTLAS 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24426594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents_AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep.*
6: /cgn2_6/ptodata/2/iaa/backfiles2.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	44	100.0	9	1	US-08-619-645-1	Sequence 1, Appli
2	44	100.0	9	2	US-08-664-495-1	Sequence 1, Appli
3	44	100.0	18	1	US-08-619-645-5	Sequence 5, Appli
4	44	100.0	18	2	US-08-664-495-5	Sequence 5, Appli
5	44	100.0	18	4	US-09-171-265-40	Sequence 40, Appl
6	43	65.2	478	3	US-08-870-198A-14	Sequence 14, Appl
7	43	65.2	478	4	US-08-870-198A-15	Sequence 15, Appl
8	43	65.2	687	2	US-08-664-495-6	Sequence 6, Appli
9	43	65.2	687	3	US-08-619-645-6	Sequence 6, Appli
10	43	65.2	687	4	US-09-043-623-6	Sequence 6, Appli
11	43	65.2	846	2	US-08-489-819-4	Sequence 4, Appli
12	43	65.2	846	3	US-09-151-011-4	Sequence 4, Appli
13	43	65.2	846	4	US-09-039-198A-2	Sequence 2, Appli
14	43	65.2	846	4	US-09-039-198A-4	Sequence 4, Appli
15	43	65.2	846	4	US-09-043-623-4	Sequence 4, Appli
16	38	86.4	18	4	US-09-171-265-41	Sequence 41, Appl
17	36	81.8	63	1	US-07-612-624-12	Sequence 12, Appl
18	36	81.8	270	1	US-07-612-624-5	Sequence 5, Appli
19	36	81.8	270	1	US-07-612-624-8	Sequence 8, Appli
20	36	81.8	68	2	US-08-664-915-2	Sequence 2, Appli
21	36	81.8	414	2	US-08-664-915-4	Sequence 4, Appli
22	35	79.5	551	2	US-08-724-051-2	Sequence 2, Appli
23	35	79.5	551	4	US-09-052-778-16	Sequence 16, Appl
24	32	72.7	18	1	US-08-619-645-6	Sequence 6, Appli
25	32	72.7	18	2	US-08-664-495-6	Sequence 6, Appli
26	31	70.5	198	4	US-09-080-983-15	Sequence 15, Appl
27	28	63.6	134	4	US-09-170-147A-89	Sequence 89, Appl
28	28	63.6	111	1	US-07-942-245-27	Sequence 27, Appl
29	28	63.6	151	2	US-08-637-759B-449	Sequence 449, Appl
30	28	63.6	151	3	US-08-871-455A-449	Sequence 449, Appl
31	28	63.6	151	4	US-09-201-945-449	Sequence 449, Appl
32	28	63.6	158	4	US-09-215-221-22	Sequence 22, Appl
33	28	63.6	183	4	US-09-059-769-6	Sequence 6, Appl
34	28	63.6	345	1	US-08-154-570-14	Sequence 14, Appl
35	28	63.6	384	2	US-08-637-759B-475	Sequence 475, Appl
36	28	63.6	384	3	US-08-871-455A-475	Sequence 475, Appl
37	28	63.6	384	4	US-09-201-945-475	Sequence 475, Appl
38	28	63.6	418	2	US-08-494-907-18	Sequence 18, Appl
39	28	63.6	418	5	PCT-US96-10986-18	Sequence 18, Appl
40	28	63.6	533	1	US-08-488-805A-6	Sequence 6, Appli
41	28	63.6	533	4	US-09-385-259-2	Sequence 2, Appli
42	28	63.6	715	4	US-09-215-221-25	Sequence 25, Appl
43	28	63.6	727	1	US-08-424-424B-2	Sequence 2, Appli
44	28	63.6	727	5	PCT-US94-05363A-2	Sequence 2, Appli
45	28	63.6	1050	4	US-09-428-711A-16	Sequence 16, Appl

ALIGNMENTS

RESULT 1
US-08-619-645-1
; Sequence 1, Application US/08619645
; Patent No. 5736507
; GENERAL INFORMATION:
; APPLICANT: Roots, Anna M.B.
; APPLICANT: Verheijden, Gijsbertus F.M.
; TITLE OF INVENTION: No. 5736507el peptides derived from
; TITLE OF INVENTION: autoantigen for use in immunotherapy of autoimmune
; TITLE OF INVENTION: disease
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Ah & No. 5736507el Patent Department
; STREET: 1300 Piccard Drive, Suite 206
; CITY: Rockville
; STATE: Maryland
; COUNTRY: U.S.A.
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25(EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/619,645
; FILING DATE: 25-MAR-1996
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Gormley, Mary E.
; REGISTRATION NUMBER: 34,409
; REFERENCE/DOCKET NUMBER: 1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 947-4433
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-619-645-1

Query Match 100.0%; Score 44; DB 1; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.7e-06;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FGRSFTLAS 9
|||||||||
Db 1 FGRSFTLAS 9

```
RESULT   2
US-08-634-493-1
; Sequence 1, Application US/08634493
; Patent No. 5843449
;  GENERAL INFORMATION:
;     APPLICANT:  A.M.H. Boots
;     APPLICANT:  G.F.M. Verheijden
;     APPLICANT:  E.S. Bos
;     TITLE OF INVENTION:  No. 5843449cl Peptides derived from autoantigens for use
;     TITLE OF INVENTION:  in Immunotherapy of Autoimmune Diseases
;     NUMBER OF SEQUENCES:  10
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Akzo No. 5843449cl Patent Department
;       STREET:  1300 Piccard Drive, Suite 206
;       CITY:  Rockville
;       STATE:  Maryland
;       COUNTRY:  USA
;       ZIP:  20850
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/634,493
;       FILING DATE:  18-APR-1996
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/619,645
;       FILING DATE:  25-MAR-1996
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  PCT/EP95/04201
;       FILING DATE:  25-OCT-1995
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  NL 9420l287
;       FILING DATE:  27-OCT-1994
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  NL 952008860
;       FILING DATE:  07-APR-1995
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (30l) 258-5200
;       TELEFAX:  (30l) 977-0847
;   INFORMATION FOR SEQ ID NO:  l:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  9 amino acids
;       TYPE:  amino acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  peptide
US-08-634-493-1


  Query Match             100.0%;  Score 44;  DB 2;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 1.7e-05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 FGRSFTLAS 9
           |||||||||
Db       1 FGRSFTLAS 9


RESULT   3
US-08-619-645-5
; Sequence 5, Application US/08619645
; Patent No. 5746507
;  GENERAL INFORMATION:
;     APPLICANT:  Boots, Anna M.H.
;     APPLICANT:  Verheijden, Gijsbertus F.M.
;     TITLE OF INVENTION:  No. 5746507cl peptides derived from
;     TITLE OF INVENTION:  autoantigen for use in Immunotherapy of Autoimmune
;     TITLE OF INVENTION:  disease
;     NUMBER OF SEQUENCES:  9
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Akzo No. 5746507cl Patent Department
;       STREET:  1300 Piccard Drive, Suite 206
;       CITY:  Rockville
;       STATE:  Maryland
;       COUNTRY:  U.S.A.
;       ZIP:  20850
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25(EPO)
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/619,645
;       FILING DATE:  25-MAR-1996
;       CLASSIFICATION:  514
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Gormley, Mary E.
;       REGISTRATION NUMBER:  34,409
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (30l) 947-4433
;   INFORMATION FOR SEQ ID NO:  5:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  13 amino acids
;       TYPE:  amino acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  peptide
US-08-619-645-5


  Query Match             100.0%;  Score 44;  DB 1;  Length 13;
  Best Local Similarity   100.0%;  Pred. No. 2.9e-05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 FGRSFTLAS 9
           |||||||||
Db       3 FGRSFTLAS 11


RESULT   4
US-08-634-493-5
; Sequence 5, Application US/08634493
; Patent No. 5843449
;  GENERAL INFORMATION:
;     APPLICANT:  A.M.H. Boots
;     APPLICANT:  G.F.M. Verheijden
;     APPLICANT:  E.S. Bos
;     TITLE OF INVENTION:  No. 5843449cl Peptides derived from autoantigens for use
;     TITLE OF INVENTION:  in Immunotherapy of Autoimmune Diseases
;     NUMBER OF SEQUENCES:  10
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Akzo No. 5843449cl Patent Department
;       STREET:  1300 Piccard Drive, Suite 206
;       CITY:  Rockville
;       STATE:  Maryland
;       COUNTRY:  USA
;       ZIP:  20850
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/634,493
;       FILING DATE:  18-APR-1996
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/619,645
;       FILING DATE:  25-MAR-1996
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  PCT/EP95/04201
;       FILING DATE:  25-OCT-1995
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  NL 9420l287
```

```
;       FILING DATE:  27-APR-1994
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  NL 9320 0860
;       FILING DATE:  07-APR-1993
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  +01 25 5 520
;       TELEFAX:  (+01) 47 0847
;   INFORMATION FOR SEQ ID NO:  5:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  13 amino acids
;       TYPE:  amino acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  peptide
US-08-341 103-5


  Query Match             100.0%;  Score 44;  DB 2;  Length 13;
  Best Local Similarity   100.0%;  Pred. No. 0.00032;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 FGRSFILAS 9
           |||||||||
Db       8 FGRSFILAS 16


RESULT   5
US-09-171-705-40
; Sequence 40, Application US/09171705
; Patent No. 6184201
; GENERAL INFORMATION:
;  APPLICANT: BOS, JOHANNES D.H.
;  APPLICANT: VERHEIJDEN, GILBERTUS F.M.
;  TITLE OF INVENTION: NOVEL PEPTIDES SUITABLE FOR USE IN ANTIGEN SPECIFIC
;  TITLE OF INVENTION: IMMUNOSUPPRESSIVE THERAPY
;  FILE REFERENCE:  /96198 US
;  CURRENT APPLICATION NUMBER: US/09/171,705
;  CURRENT FILING DATE: 1999-02-09
;  NUMBER OF SEQ ID NOS: 78
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 40
;   LENGTH: 16
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: /note= "Description of Artificial Sequence: PEPTIDE FRAGM
;   OTHER INFORMATION: SEQUENCE 8 HUMAN CARTILAGE (HC)  99 AS 113
US-09-171-705-40


  Query Match             100.0%;  Score 44;  DB 4;  Length 16;
  Best Local Similarity   100.0%;  Pred. No. 0.0004;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 FGRSFILAS 9
           |||||||||
Db       8 FGRSFILAS 16


RESULT   6
US-09-039-198A-14
; Sequence 14, Application US/09039198A
; Patent No. 6200951
;  GENERAL INFORMATION:
;    APPLICANT:  Gray, Patrick W.
;    APPLICANT:  Ljolker, Larry W.
;    TITLE OF INVENTION:  CHITINASE CHITIN BINDING FRAGMENTS
;    NUMBER OF SEQUENCES:  34
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray & Borun
;      STREET:  233 South Wacker Drive/6300 Sears Tower
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  United States of America
;      ZIP:  60606-6402
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/039,198A
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Rin-Laures, Li-Hsien
;      REGISTRATION NUMBER:  33,547
;      REFERENCE/DOCKET NUMBER:  27866/34391
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (312) 474-6300
;      TELEFAX:  (312) 474-0448
;  INFORMATION FOR SEQ ID NO:  14:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  373 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-09-039-198A-14


  Query Match             93.2%;  Score 41;  DB 4;  Length 373;
  Best Local Similarity   88.9%;  Pred. No. 0.6;
  Matches    8;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 FGRSFILAS 9
           :|||||||:
Db     246 YGRSFILAS 254


RESULT   7
US-09-039-198A-15
; Sequence 15, Application US/09039198A
; Patent No. 6200951
;  GENERAL INFORMATION:
;    APPLICANT:  Gray, Patrick W.
;    APPLICANT:  Ljolker, Larry W.
;    TITLE OF INVENTION:  CHITINASE CHITIN BINDING FRAGMENTS
;    NUMBER OF SEQUENCES:  34
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray & Borun
;      STREET:  233 South Wacker Drive/6300 Sears Tower
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  United States of America
;      ZIP:  60606-6402
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/039,198A
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Rin-Laures, Li-Hsien
;      REGISTRATION NUMBER:  33,547
;      REFERENCE/DOCKET NUMBER:  27866/34391
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (312) 474-6300
;      TELEFAX:  (312) 474-0448
;  INFORMATION FOR SEQ ID NO:  15:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  373 amino acids
;      TYPE:  amino acid
```

```
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  peptide
US-09-039-198A-15


  Query Match             93.2%;  Score 41;  DB 4;  Length 47;
  Best Local Similarity   88.9%;  Pred. No. 0.6;
  Matches    8;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 FGRSFILAS 9
           :|||.|:|:
Db      246 YGRSFTLAS 254


RESULT   8
US-08-486-839-6
; Sequence 6, Application US/08486839
; Patent No. 5928928
;  GENERAL INFORMATION:
;     APPLICANT:
;     TITLE OF INVENTION:  A human chitinase, its preparation
;     TITLE OF INVENTION:  production, its use for decomposing chitin, its use
;     TITLE OF INVENTION:  in therapy or prophylaxis against infection diseases.
;     NUMBER OF SEQUENCES:  16
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Hoffmann & Baron
;       STREET:  350 Jericho Turnpike
;       CITY:  Jericho
;       STATE:  New York
;       COUNTRY:  United States of America
;       ZIP:  11758
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.30
;       SOFTWARE:
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/486,839
;       FILING DATE:  07 June 1995
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Baron, Ronald J.
;       REGISTRATION NUMBER:  29,281
;       REFERENCE/DOCKET NUMBER:  294 26
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (516) 822 3550
;       TELEFAX:  (516) 822 3582
;   INFORMATION FOR SEQ ID NO:  6:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  387 amino acids
;       TYPE:  amino acid
;       STRANDEDNESS:  unknown
;       TOPOLOGY:  unknown
;     MOLECULE TYPE:  protein
;     HYPOTHETICAL:  N
US-08-486-839-6


  Query Match             93.2%;  Score 41;  DB 2;  Length 387;
  Best Local Similarity   88.9%;  Pred. No. 0.62;
  Matches    8;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 FGRSFILAS 9
           :|||.|:|:
Db     267 YGRSFTLAS 275


RESULT   9
US-09-151-011-6
; Sequence 6, Application US/09151011
; Patent No. 6077142
;  GENERAL INFORMATION:
;     APPLICANT:
;     TITLE OF INVENTION:  A Human Chitinase, Its Preparation
;     TITLE OF INVENTION:  Production, Its Use For Decomposing Chitin, Its Use
;     TITLE OF INVENTION:  Therapy or Prophylaxis Against Infection Diseases
;     NUMBER OF SEQUENCES:  16
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Hoffmann & Baron, LLP
;       STREET:  6900 Jericho Turnpike
;       CITY:  Syosset
;       STATE:  New York
;       COUNTRY:  United States of America
;       ZIP:  11791
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC DOS/MS DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.30
;       SOFTWARE:
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  09/151,011
;       FILING DATE:  10 September 1998
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Morriss, Robert C.
;       REGISTRATION NUMBER:  42,915
;       REFERENCE/DOCKET NUMBER:  294 42 DIV
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (516) 822 3550
;       TELEFAX:  (516) 822-3582
;   INFORMATION FOR SEQ ID NO:  6:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  387 amino acids
;       TYPE:  amino acid
;       STRANDEDNESS:  unknown
;       TOPOLOGY:  unknown
;     MOLECULE TYPE:  protein
;     HYPOTHETICAL:  NO
US-09-151-011-6


  Query Match             93.2%;  Score 41;  DB 3;  Length 387;
  Best Local Similarity   88.9%;  Pred. No. 0.62;
  Matches    8;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 FGRSFILAS 9
           :|||.|:|:
Db     267 YGRSFTLAS 275


RESULT   10
US-09-343-623-6
; Sequence 6, Application US/09343623
; Patent No. 6303118
;  GENERAL INFORMATION:
;     APPLICANT:
;     TITLE OF INVENTION:  A Human chitinase, its preparation
;     TITLE OF INVENTION:  production, its use for decomposing chitin, its use
;     TITLE OF INVENTION:  in therapy or prophylaxis against infection diseases.
;     NUMBER OF SEQUENCES:  16
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Hoffmann & Baron
;       STREET:  350 Jericho Turnpike
;       CITY:  Jericho
;       STATE:  New York
;       COUNTRY:  United States of America
;       ZIP:  11758
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC DOS/MS DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.30
;       SOFTWARE:
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  09/343,623
```

```
;       FILING DATE:
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US/09/151,011
;       FILING DATE:  07-Jun-1998
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Baron, Ronald J.
;       REGISTRATION NUMBER:  29,586
;       REFERENCE/DOCKET NUMBER:  294.26
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (516) 822-4550
;       TELEFAX:  (516) 822-4582
;   INFORMATION FOR SEQ ID NO:  5:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  387 amino acids
;       TYPE:  amino acid
;       STRANDEDNESS:  unknown
;       TOPOLOGY:  unknown
;     MOLECULE TYPE:  protein
;     HYPOTHETICAL:  NO
US-09-744-282-1

  Query Match             93.2%;  Score 41;  DB 4;  Length 387;
  Best Local Similarity   88.9%;  Pred. No. 0.62;
  Matches    8;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 FGRSFTLAS 9
             :|||||||||
Db     267 YGRSFTLAS 275


RESULT  11
US-08-486-839-4
; Sequence 4, Application US/08486839
; Patent No. 5928928
;  GENERAL INFORMATION:
;    APPLICANT:
;    TITLE OF INVENTION:  A human chitinase, its recombinant
;    TITLE OF INVENTION:  production, its use for decomposing chitin, its use
;    TITLE OF INVENTION:  in therapy or prophylaxis against infection diseases.
;    NUMBER OF SEQUENCES:  16
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Hoffmann & Baron
;      STREET:  350 Jericho Turnpike
;      CITY:  Jericho
;      STATE:  New York
;      COUNTRY:  United States of America
;      ZIP:  11753
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;      SOFTWARE:
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/486,839
;      FILING DATE:  07-June-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Baron, Ronald J.
;      REGISTRATION NUMBER:  29,586
;      REFERENCE/DOCKET NUMBER:  294.29
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  516 822 4550
;      TELEFAX:  (516) 822 4582
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  466 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  unknown
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:  protein
;    HYPOTHETICAL:  NO
US-08-486-839-4

  Query Match             93.2%;  Score 41;  DB 4;  Length 466;
  Best Local Similarity   88.9%;  Pred. No. 0.77;
  Matches    8;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 FGRSFTLAS 9
             :|||||||||
Db     267 YGRSFTLAS 275


RESULT  12
US-09-151-011-4
; Sequence 4, Application US/09151011
; Patent No. 6057142
;  GENERAL INFORMATION:
;    APPLICANT:
;    TITLE OF INVENTION:  A Human Chitinase, Its Recombinant
;    TITLE OF INVENTION:  Production, Its Use For Decomposing Chitin, Its Use In
;    TITLE OF INVENTION:  Therapy or Prophylaxis Against Infection Diseases.
;    NUMBER OF SEQUENCES:  16
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Hoffmann & Baron, LLP
;      STREET:  6900 Jericho Turnpike
;      CITY:  Syosset
;      STATE:  New York
;      COUNTRY:  United States of America
;      ZIP:  11791
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;      SOFTWARE:
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/151,011
;      FILING DATE:  10-September-1998
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Morriss, Robert C.
;      REGISTRATION NUMBER:  42,910
;      REFERENCE/DOCKET NUMBER:  294.32 DIV
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (516) 822-4550
;      TELEFAX:  (516) 822-4582
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  466 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  unknown
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:  protein
;    HYPOTHETICAL:  NO
US-09-151-011-4

  Query Match             93.2%;  Score 41;  DB 3;  Length 466;
  Best Local Similarity   88.9%;  Pred. No. 0.77;
  Matches    8;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 FGRSFTLAS 9
             :|||||||||
Db     267 YGRSFTLAS 275


RESULT  13
US-09-039-198A-2
; Sequence 2, Application US/09039198A
; Patent No. 6200951
;  GENERAL INFORMATION:
;    APPLICANT:  Gray, Patrick W.
;    APPLICANT:  Ljoerker, Larry W.
;    TITLE OF INVENTION:  CHITINASE CHITIN BINDING FRAGMENTS
;    NUMBER OF SEQUENCES:  84
```

```
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray & Borun
;       STREET:  233 South Wacker Drive/6300 Sears Tower
;       CITY:  Chicago
;       STATE:  Illinois
;       COUNTRY:  United States of America
;       ZIP:  60606-6402
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC DOS/MS DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.30
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/09/039,198A
;       FILING DATE:
;       CLASSIFICATION:  435
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Rin Lauren, Li Hsien
;       REGISTRATION NUMBER:  33,547
;       REFERENCE/DOCKET NUMBER:  27866/44391
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (312) 474-6300
;       TELEFAX:  (312) 474-0448
;   INFORMATION FOR SEQ ID NO:  2:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  466 amino acids
;       TYPE:  amino acid
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  protein
US-09-039-198A-2


  Query Match             93.2%;  Score 41;  DB 4;  Length 466;
  Best Local Similarity   88.9%;  Pred. No. 0.77;
  Matches    8;  Conservative    1;  Mismatches    0;  Indels    1;  Gaps    0;

Qy       1 FGRSFILAS 9
             :|||||:|
Db     267 YGKSFILAS 275


RESULT  14
US-09-039-198A-4
; Sequence 4, Application US/09039198A
; Patent No. 6200565
; GENERAL INFORMATION:
;     APPLICANT:  Gray, Patrick W.
;     APPLICANT:  Tjoelker, Larry W.
;     TITLE OF INVENTION:  CHIMERIC CHITIN BINDING FRAGMENTS
;     NUMBER OF SEQUENCES:  34
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray & Borun
;       STREET:  233 South Wacker Drive/6300 Sears Tower
;       CITY:  Chicago
;       STATE:  Illinois
;       COUNTRY:  United States of America
;       ZIP:  60606-6402
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC DOS/MS DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.30
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/09/039,198A
;       FILING DATE:
;       CLASSIFICATION:  435
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Rin Lauren, Li Hsien
;       REGISTRATION NUMBER:  33,547
;       REFERENCE/DOCKET NUMBER:  27866/44391
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (312) 474-6300
;       TELEFAX:  (312) 474-0448
;   INFORMATION FOR SEQ ID NO:  4:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  466 amino acids
;       TYPE:  amino acid
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  protein
US-09-039-198A-4


  Query Match             93.2%;  Score 41;  DB 4;  Length 466;
  Best Local Similarity   88.9%;  Pred. No. 0.77;
  Matches    8;  Conservative    1;  Mismatches    0;  Indels    1;  Gaps    0;

Qy       1 FGRSFILAS 9
             :|||||:|
Db     267 YGKSFILAS 275


RESULT  15
US-09-343-624-4
; Sequence 4, Application US/09343624
; Patent No. 6303138
; GENERAL INFORMATION:
;     APPLICANT:
;     TITLE OF INVENTION:  A human chitinase, its recombinant
;     TITLE OF INVENTION:  production, its use for dosing [illegible]
;     TITLE OF INVENTION:  in therapy or prophylaxis against [illegible]
;     NUMBER OF SEQUENCES:  16
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Hoffmann & Baron
;       STREET:  45 Jericho Turnpike
;       CITY:  Jericho
;       STATE:  New York
;       COUNTRY:  United States of America
;       ZIP:  11753
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC DOS/MS DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.30
;       SOFTWARE:
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/09/343,624
;       FILING DATE:
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/486,849
;       FILING DATE:  07 June 1995
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Baron, Ronald J.
;       REGISTRATION NUMBER:  29,281
;       REFERENCE/DOCKET NUMBER:  294.29
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (516) 822-3550
;       TELEFAX:  (516) 822-3582
;   INFORMATION FOR SEQ ID NO:  4:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  466 amino acids
;       TYPE:  amino acid
;       STRANDEDNESS:  Unknown
;       TOPOLOGY:  Unknown
;     MOLECULE TYPE:  protein
;     HYPOTHETICAL:  No
US-09-343-624-4


  Query Match             93.2%;  Score 41;  DB 4;  Length 466;
  Best Local Similarity   88.9%;  Pred. No. 0.77;
  Matches    8;  Conservative    1;  Mismatches    0;  Indels    1;  Gaps    0;

Qy       1 FGRSFILAS 9
             :|||||:|
Db     267 YGKSFILAS 275
```

Search completed: August 26, 2002, 06:23:25
Job time: 8.0 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 26, 2002, 06:12:44 ; Search time 7.5 Seconds
(without alignments)
10,743 Million cell updates/sec

Title: US-09-744-282-2
Perfect score: 43
Sequence: 1 FTLASSEIG 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 99086434 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length:
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: PIR 77:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	43	100.0	483	2	A49562	cartilage glycopro
2	37	86.0	441	2	I31090	probable ATP bindi
3	33	76.7	385	2	S61727	heparin binding gl
4	32	74.4	251	1	C69343	conserved hypothet
5	32	74.4	261	2	F64235	hypothetical prote
6	32	74.4	360	2	S66117	yacK protein - Bac
7	32	74.4	405	2	S81581	breast regressing
8	32	74.4	573	2	A47781	hypothetical prote
9	32	74.4	989	2	A62209	two component sens
10	31	72.1	187	2	S69542	hypothetical prote
11	31	72.1	201	2	G87355	hypothetical prote
12	31	72.1	318	2	E70178	hypothetical prote
13	31	72.1	320	2	I24905	hypothetical prote
14	31	72.1	324	2	T20681	hypothetical prote
15	31	72.1	327	2	D83431	muconate cycloisom
16	31	72.1	413	2	I43867	hypothetical prote
17	31	72.1	445	2	B40990	flagellar distal c
18	31	72.1	450	1	P2W106	L2 protein - human
19	31	72.1	498	2	I20291	hypothetical prote
20	31	72.1	515	2	S38270	amidase (EC 3.5.1.
21	31	72.1	590	1	GBBP14	baseplate protein
22	31	72.1	605	2	E69173	cadmium efflux ATP
23	31	72.1	868	2	I49420	hypothetical prote
24	31	72.1	961	2	I49009	hypothetical prote
25	31	72.1	977	2	E14152	synaptic scaffoldi
26	30	69.8	167	2	AB2706	acetyltransferase
27	30	69.8	167	2	B97576	hypothetical prote
28	30	69.8	241	2	S71434	endo-1,4-beta-xyla
29	30	69.8	246	2	A25729	retrovirus-related
30	30	69.8	266	2	G70788	hypothetical pr
31	30	69.8	317	2	S45708	MSB receptor - bo
32	30	69.8	346	2	E72672	hypothetical pr
33	30	69.8	365	2	AG1246	amitotriptyl'ase
34	30	69.8	375	2	AH0306	probable exported
35	30	69.8	433	2	S15471	pol polyprotein
36	30	69.8	447	2	I06004	1-aminocyclopropan
37	30	69.8	461	2	I03414	probable L-amino-o
38	30	69.8	488	2	D70876	probable polyketid
39	30	69.8	596	2	AC1418	two-component sens
40	30	69.8	782	2	I48746	semaphorin C - mou
41	30	69.8	893	2	I03864	hypothetical prote
42	29	67.4	163	2	S11149	amide protein - Sti
43	29	67.4	167	2	AG2292	hypothetical prote
44	29	67.4	178	2	C40956	interleukin-1 rece
45	29	67.4	197	2	T43066	hypothetical prote

ALIGNMENTS

RESULT 1
A49562
cartilage glycoprotein gp39 precursor - human
N;Alternate names: 39K synovial protein
C;Species: Homo sapiens (man)
C;Date: 23-Mar-1995 #sequence_revision 23-Mar-1995 #text_change 22-Jun-1999
C;Accession: A49562; S10677; A43162
R;Hakala, B.E.; White, C.; Recklies, A.D.
J. Biol. Chem. 268, 25803-25810, 1993
A;Title: Human cartilage gp-39, a major secretory product of articular chondrocytes a
A;Reference number: A49562; MUID:94064658
A;Accession: A49562
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-483 <HAK>
A;Cross-references: GB:M80927; NID:g348911; PIDN:AAA16074.1; PID:g349912
R;Nyirkos, P.; Golds, E.E.
Biochem. J. 269, 265-268, 1990
A;Title: Human synovial cells secrete a 39 kDa protein similar to a bovine mammary pr
A;Reference number: S10677; MUID:90328083
A;Accession: S10677
A;Molecule type: protein
A;Residues: 22-40,'X',42-45 <NYZ>
C;Superfamily: Streptomyces chitinase chi40
C;Keywords: cartilage; extracellular protein; glycoprotein
F;1-21/Domain: signal sequence #status predicted <SIG>
F;22-483/Product: cartilage glycoprotein gp39 #status predicted <MAT>

Query Match 100.0%; Score 43; DB 2; Length 483;
Best Local Similarity 100.0%; Prod. No. 0.11;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FTLASSEIG 9
|||||||||
Db 265 FTLASSEIG 273

RESULT 2
I31090
probable ATP-binding protein - Serratia marcescens
C;Species: Serratia marcescens
C;Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 17-Mar-2000
C;Accession: I31090
R;Saini, P.; Climent, N.; Fiquer, N.; Sanchez, C.; Merino, S.; Rubires, X.; Aguilar, A
J. Bacteriol. 181, 1883-1891, 1999
A;Title: Genetic analysis of the Serratia marcescens N28b O6 antigen gene cluster.
A;Reference number: 220974; MUID:99174913
A;Accession: I31090
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-441 <SAI>

A;Cross references: EMBL:AE048816; NID:g2828669; PID:g2828672; PIDN:AAC00182.1
C;Genetics:
A;Gene: wzt
C;Superfamily: unassigned ATP-binding cassette proteins; ATP binding cassette homology

Query Match 86.0%; Score 37; DB 2; Length 441;
Best Local Similarity 100.0%; Pred. No. 2.9;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FTLASSEI 8
||||||||
Db 332 FTLASSEI 339

RESULT 3
S51327
heparin-binding glycoprotein 38K - pig
C;Species: Sus scrota domestica (domestic pig)
C;Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 22-Jun-1999
C;Accession: S51327
R;Shackelton, L.M.; Mann, D.M.; Millis, A.J.T.
submitted to the EMBL Data Library, January 1995
A;Description: Identification of a 38kDa heparin-binding glycoprotein (gp38k) in differe
A;Reference number: S51327
A;Accession: S51327
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-383 <SHA>
A;Cross-references: EMBL:Z47803; NID:g634097; PIDN:CAA87764.1; PID:g634098
C;Superfamily: Streptomyces chitinase chi40

Query Match 76.7%; Score 33; DB 2; Length 383;
Best Local Similarity 87.5%; Pred. No. 20;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FTLASSEI 8
:|:|:|:|
Db 265 FTLASSKI 272

RESULT 4
C69046
conserved hypothetical protein MTH1349 - Methanobacterium thermoautotrophicum (strain De
C;Species: Methanobacterium thermoautotrophicum
C;Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 22-Oct-1999
C;Accession: C69046
R;Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.;
; Qin, D.; Spadafora, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiwani, N.
ki, S.-; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Nelissen, J.; Reeve, J.N.
J. Bacteriol. 179, 7135-7155, 1997
A;Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: funct
A;Reference number: A69030; MUID:98037514
A;Accession: C69046
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-251 <MTH>
A;Cross-references: GB:AE000898; GB:AE000666; NID:g2622495; PIDN:AAB85820.1; PID:g2622495
A;Experimental source: strain Delta H
C;Genetics:
A;Gene: MTH1349

Query Match 74.4%; Score 32; DB 2; Length 251;
Best Local Similarity 66.7%; Pred. No. 21;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FTLASSEIG 9
||:|:|:|:|
Db 7 FTIFSTEIG 15

RESULT 5
P84325
hypothetical protein VNG1735C [imported] - Halobacterium sp. NRC-1
C;Species: Halobacterium sp. NRC-1
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 16-Jul-2001
C;Accession: P84325
R;Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky
; Leithauser, B.; Keller, K.; Cruz, R.; Dawson, M.J.; Booth, J.W.; Marti-Renom, J.C.; Ba
Jung, K.H.; Alam, M.; Freitas, L.
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
A;Authors: Ng, W.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.;
A;Title: Genome sequence of Halobacterium species NRC-1.
A;Reference number: A84160; MUID:21054483
A;Accession: P84325
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-261 <SFD>
A;Cross-references: GB:AE004437; NID:g10581197; PIDN:AAG19970.1; GSPDB:GN00108
C;Genetics:
A;Gene: VNG1735C

Query Match 74.4%; Score 32; DB 2; Length 261;
Best Local Similarity 66.7%; Pred. No. 22;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 FTLASSEIG 9
||| :|:||
Db 87 FTLVESDTG 95

RESULT 6
S66117
yacK protein - Bacillus subtilis
C;Species: Bacillus subtilis
C;Date: 28-Oct-1996 #sequence_revision 15-Mar-1997 #text_change 21-Jul-2000
C;Accession: S66117; C69741
R;Ogasawara, N.; Nakai, S.; Yoshikawa, H.
DNA Res. 1, 1-14, 1994
A;Title: Systematic sequencing of the 180 kilobase region of the Bacillus subtilis che
A;Reference number: S63967; MUID:96051485
A;Accession: S66117
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-360 <OGA>
A;Cross-references: EMBL:D26185; NID:g467520; PIDN:BAA05232.1; PID:g467521
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, October 1993
R;Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Ber
C.; Bron, S.; Bruschi, C.; Caldwell, B.; Capuano, V.; Carter, N.M.;
A., Ehrlich, S.D., Emmerson, P.T., Entian, K.D.; Errington, J.; Fabret, C.; Ferrari,
Nature 390, 249-256, 1997
A;Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Ga
Tech, J.; Harwood, C.R.; Hennat, A.; Hilbert, H.; Holsappel, S.; Hosoi,
Koetter, P.; Konietzny, U.; Kreft, S.; Krogh, A.; Kurita, K.; Lapidus,
A;Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Ma
y, M.; Ogawa, K.; Ogiwara, A.; Oguro, A.; Ono, H.; Park, S.H.; Parro,
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie,
A;Authors: Schleich, S.; Schroeter, R.; Scottoni, E.; Sekiguchi,
Takeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni,
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto,
A;Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin,
A;Title: The complete genome sequence of the Gram-positive bacterium
A;Reference number: A59780; MUID:98044033
A;Accession: C69741
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-360 <KUN>
A;Cross-references: GB:Z99104; GB:AL009126; NID:g2632267; PIDN:CAB15904.1; PID:g2632268
A;Experimental source: strain 168
C;Genetics:
A;Gene: yacK

Query Match 74.4%; Score 32; DB 2; Length 566;
Best Local Similarity 75.0%; Pred. No. 42;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 2 TLASSETG 9
:|:|:|.|
Db 131 TISSSEIG 108

RESULT 7
S61951
breast repressing protein brp39 precursor - mouse
C;Species: Mus musculus (house mouse)
C;Date: 19-Mar-1997 #sequence_revision 25-Apr-1997 #text_change 20-Jun-2000
C;Accession: S61951; S61956; I14171
R;Morrison, B.W.; Leder, P.
Oncogene 9, 3417-3426, 1994
A;Title: Novel acidic and basic 39 kDa/pI 5.6 mammary tumors that share genetic markers generally
A;Reference number: I48171; MUID:95067797
A;Accession: S61951
A;Status: nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1-465 <MOR>
A;Cross-references: EMBL:X96035; NID:g1035896; PIDN:CAA65065.1; PID:g1067666
R;Morrison, B.W.
submitted to the EMBL Data Library, November 1995
A;Reference number: S61950
A;Accession: S61950
A;Molecule type: mRNA
A;Residues: 1-245,'C',247-330,'E',332-350,'MVWALLALLGFQGTQAEKRTFTKLNAIKLALA' <MOR2>
A;Cross-references: EMBL:X96035; NID:g1035896; PIDN:CAA65065.1; PID:g1067666
A;Note: the differences at the carboxyl end are due to a frameshift error
C;Genetics:
A;Gene: brp39
C;Superfamily: Streptomyces chitinase chi40
F;1-21/Domain: signal sequence #status predicted <SIG>
F;22-405/Product: breast repressing protein brp39 #status predicted <MAT>

Query Match 74.4%; Score 32; DB 2; Length 405;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FTLASSE 7
|:|:|.|
Db 266 FTLASSE 272

RESULT 8
AB4782
hypothetical protein At2g36560 [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse ear cress)
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
C;Accession: AB4782
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Kotani, H.; Matta, E.; Nierman, W.C.; Feldblyum, T.; Hansen, H.; Adams, M.; Sodergren, E.J.;
ouse, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
Nature 402, 761-768, 1999
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A;Reference number: AB4120; MUID:20083487
A;Accession: AB4782
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-574 <LIN>
A;Cross-references: GB:AE002093; NID:g4581154; PIDN:AAD24638.1; GSPDB:GN00139
C;Genetics:
A;Gene: At2g36560
A;Map position: 2

Query Match 74.4%; Score 32; DB 2; Length 574;
Best Local Similarity 87.5%; Pred. No. 54;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 TLASSETG 9
|||||.||
Db 290 TLASSSTG 297

RESULT 9
AB2209
two-component sensor histidine kinase alr4225 [imported] - Anabaena sp. (strain PCC 7
C;Species: Anabaena sp.
A;Note: Anabaena sp. (strain PCC 7120) is a synonym of Nostoc sp. strain PCC 7120
C;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 11-Jan-2002
C;Accession: AB2209
R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguc
Nakazaki, N.; Shimpo, S.; Obaishi, M.; Iwabuchi, M.; Ogawa, M.; Yasuda, M.; Tabata
DNA Res. 8, 205-213, 2001
A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium
A;Reference number: AB1807; MUID:21595286; PMID:11759984
A;Accession: AB2209
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-986 <KUK>
A;Cross-references: GB:BA000019; PIDN:BAB74924.1; PID:g17132
C;Genetics:
A;Gene: alr4225

Query Match 74.4%; Score 32; DB 2; Length 986;
Best Local Similarity 66.7%; Pred. No. 99;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FTLASSETG 9
|:|:|||||
Db 536 FTVKNSETG 544

RESULT 10
S69542
hypothetical protein 84 - phage HP1
C;Species: phage HP1
C;Date: 06-Dec-1996 #sequence_revision 14-Mar-1997 #text_change 08-Oct-1999
C;Accession: S69542
R;Esposito, D.; Fitzmaurice, W.P.; Benjamin, R.C.; Goodman, S.I.; Waldman, A.S.; Scoc
Nucleic Acids Res. 24, 2360-2368, 1996
A;Title: The complete nucleotide sequence of bacteriophage HP1 DNA.
A;Reference number: S69503; MUID:96279738
A;Accession: S69542
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-187 <ESP>
A;Cross-references: EMBL:U24159; NID:gU46235; PIDN:AAB09221.1; PID:gU40253
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, Apr-11 1995

Query Match 72.1%; Score 31; DB 2; Length 287;
Best Local Similarity 66.7%; Pred. No. 26;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 FTLASSETG 9
|.||:|||
Db 72 FNLAEAEIG 80

RESULT 11
G87353
hypothetical protein CC0842 [imported] - Caulobacter crescentus
C;Species: Caulobacter crescentus
C;Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 20-Apr-2001
C;Accession: G87353
R;Nierman, W.C., Feldblyum, T.V., Paulsen, I.T., Nelson, K.E., Eisen, J., Heidelberg,
B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Ko
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L., Shapiro, L.; Venter, J.C.; Fraser, C

Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A;Title: Complete Genome Sequence of Caulobacter crescentus.
A;Reference number: A87249; MUID:21173698; PMID:11259647
A;Accession: G87353
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-202 <SIE>
A;Cross-references: GB:AE005673; NID:g13422095; PIDN:AAK22827.1; GSPDB:GN00118
C;Genetics:
A;Gene: CC0084,

Query Match 72.1%; Score 41; DB 2; Length 202;
Best Local Similarity 66.7%; Pred. No. 28;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 FILASSEIG 9
:|||.|.||
Db 141 FILVSGKIG 149

RESULT 12
I20178
hypothetical protein C5B4.6 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #Sequence revision 15-Oct-1999 #text_change 15-Oct-1999
C;Accession: I20178
R;Berks, M.
submitted to the EMBL Data Library, December 1995
A;Reference number: Z19283
A;Accession: I20178
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-318 <WIL>
A;Cross-references: EMBL:Z68217; PIDN:CAA92454.1; GSPDB:GN00022; CESP:C5B4.6
A;Experimental source: clone C5B4
C;Genetics:
A;Gene: CESP:C5B4.6
A;Map position: 4
A;Introns: 25/2; 83/2; 97/3; 141/3; 208/2; 271/2

Query Match 72.1%; Score 41; DB 2; Length 318;
Best Local Similarity 77.8%; Pred. No. 47;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 FILASSEIG 9
|||||.|.|
Db 133 FILASYEFG 141

RESULT 13
T24905
hypothetical protein T14C1.1 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence revision 15-Oct-1999 #text_change 19-Feb-2002
C;Accession: T24905
R;Barlow, K.
submitted to the EMBL Data Library, April 1996
A;Reference number: Z19952
A;Accession: T24905
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-320 <WIL>
A;Cross-references: EMBL:Z70687; NID:g1256502; PIDN:CAA94616.1; GSPDB:GN00028; CESP:T14C
A;Experimental source: clone T14C1
C;Genetics:
A;Gene: CESP:T14C1.1
A;Map position: X
A;Introns: 75/2; 131/2; 199/3; 273/2;

Query Match 72.1%; Score 41; DB 2; Length 320;
Best Local Similarity 66.7%; Pred. No. 47;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FILASSEIG 9
|:||.|.||
Db 10 FSLACSEIG 18

RESULT 14
I20681
hypothetical protein F10A3.6 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #Sequence revision 15-Oct-1999 #text_change 15-Oct-1999
C;Accession: I20681
R;Lloyd, C.
submitted to the EMBL Data Library, March 1997
A;Reference number: Z19309
A;Accession: I20681
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-324 <WIL>
A;Cross-references: EMBL:Z92826; PIDN:CAB07544.1; GSPDB:GN [illegible]
A;Experimental source: clone F10A3
C;Genetics:
A;Gene: CESP:F10A3.6
A;Map position: 5
A;Introns: 64/3; 94/,

Query Match 72.1%; Score 41; DB 2; Length 324;
Best Local Similarity 75.0%; Pred. No. 48;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 ILASSEIG 9
|:||||
Db 235 FIGSSEFG 242

RESULT 15
D83331
muconate cycloisomerase I PA2509 [imported] - Pseudomonas aeruginosa (strain PA [illegible]
C;Species: Pseudomonas aeruginosa
C;Date: 15-Sep-2000 #sequence revision 15-Sep-2000 #text_change [illegible]
C;Accession: D83331
R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrener, P. [illegible]
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas [illegible]
.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A;Title: Complete genome sequence of Pseudomonas aeruginosa PA [illegible]
A;Reference number: A82950; MUID:20437337
A;Accession: D83331
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-373 <SI>
A;Cross-references: GB:AE004678; GB:AE004091; NID:g9949579 [illegible]
A;Experimental source: strain PAO1
C;Genetics:
A;Gene: catB; PA2509
C;Superfamily: muconate cycloisomerase

Query Match 72.1%; Score 41; DB 2; Length [illegible]
Best Local Similarity 66.7%; Pred. No. [illegible]
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FILASSEIG 9
:|||.|.||
Db 140 WILASGDIG 148

Search completed: August 26, 2002, 06:10:45
Job time: 106 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 26, 2002, 09:27:00 ; Search time 35.00 Seconds
(without alignments)
9.683 Million cell updates/sec

Title: US-09-744-282-2
Perfect score: 44
Sequence: 1 EILASSEIG 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105124 seqs, 46759658 residues

Total number of hits satisfying chosen parameters: 105124

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: SwissProt_40:*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	44	100.0	383	1	CH3L1_HUMAN	P36222 homo sapien
2	34	79.1	402	1	CH3L2_HUMAN	Q15782 homo sapien
3	32	74.4	360	1	YAJK_BACSU	P37573 bacillus su
4	32	74.4	382	1	CH3L1_MOUSE	P29318 mus musculu
5	32	72.1	187	1	YO44_BPHP1	P51758 bacteriopha
6	31	72.1	268	1	INBA_MYCAV	O07400 mycobacteri
7	31	72.1	402	1	CN78_SCHPO	Q09722 schizosacch
8	31	72.1	445	1	FLDL_VIBPA	Q03475 vibrio para
9	31	72.1	459	1	VL2_HPV6A	Q84207 human papil
10	31	72.1	459	1	VL2_HPV6B	P36106 human papil
11	31	72.1	598	1	VG29_BPI4	P33347 bacteriopha
12	30	69.8	246	1	FLIL_CHICK	P13994 gallus gall
13	30	69.8	417	1	MGHR_BOVIN	P32763 bos taurus
14	30	69.8	782	1	SMAB_MOUSE	Q62179 mus musculu
15	29	67.4	139	1	STRB_PSESM	Q72562 pseudomonas
16	29	67.4	178	1	ILIX_RAT	P25096 rattus norv
17	29	67.4	419	1	SSIA_ECOLI	P24235 escherichia
18	29	67.4	442	1	YEA4_YEAST	P40004 saccharomyc
19	29	67.4	444	1	ASTA_ECOLI	P76218 escherichia
20	29	67.4	448	1	YOPB_YEREN	P37741 yersinia en
21	29	67.4	481	1	YOPB_YERPS	Q06114 yersinia ps
22	29	67.4	486	1	YAM8_SCHPO	Q10009 schizosacch
23	29	67.4	573	1	RGCN_TREPA	O83456 treponema p
24	29	67.4	599	1	VG29_BPPZA	P07534 bacteriopha
25	29	67.4	642	1	FLIP_CAMJE	Q9PLW6 campylobact
26	29	67.4	659	1	AMIA_STRFN	P18791 streptococc
27	29	67.4	920	1	YML1_YEAST	Q03758 saccharomyc
28	29	67.4	1076	1	NUP1_YEAST	P20676 saccharomyc
29	29	67.4	2488	1	SPCP_RAT	Q9QWN9 rattus norv
30	28	65.1	104	1	YHK6_YEAST	P38158 saccharomyc
31	28	65.1	113	1	YO71_MYCTU	Q11039 mycobacteri
32	28	65.1	129	1	Y4PG_RHISN	P55445 rhizobium s
33	28	65.1	132	1	YGXK_BACSU	P40761 bacillus su
34	28	65.1	218	1	YU27_SIRCO	Q91KM5 sirolimus g
35	28	65.1	234	1	TRAR_AGRVI	P33989 agrobacteri
36	28	65.1	285	1	CYJ1_RHOSH	Q02769 rhodobacter
37	28	65.1	317	1	MSHR_CANFA	Q27616 canis famil
38	28	65.1	317	1	MSHR_VULVU	Q29154 vulpes vulp
39	28	65.1	321	1	TSP_BACCR	P29149 bacillus cl
40	28	65.1	322	1	GVPL_HALME	Q02237 halobacteri
41	28	65.1	339	1	RSPB_ECOLI	P38705 escherichia
42	28	65.1	339	1	Y923_TREPA	O83895 treponema p
43	28	65.1	367	1	MTC1_CHVIR	P36216 chlorella v
44	28	65.1	434	1	PYRC_THEAC	Q9HIA6 thermoplasm
45	28	65.1	417	1	CZCC_ALCEU	P38509 alcaligenes

ALIGNMENTS

RESULT 1
CH3L1_HUMAN
ID CH3L1_HUMAN STANDARD; PRT; 383 AA.
AC P36222; P80923;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Chitinase-3 like protein 1 precursor (Cartilage glycoprotein-39) (GP-
DE 39) (39 kDa synovial protein) (YKL-40).
GN CHI3L1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC TISSUE=Cartilage;
RX MEDLINE=94064658; PubMed=8245017;
RA Hakala B.E., White C., Recklies A.D.;
RT "Human cartilage gp-39, a major secretory product of articular
RT chondrocytes and synovial cells, is a mammalian member of a chitinase
RT protein family.";
RL J. Biol. Chem. 268:25803-25810(1993).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Blood;
RX MEDLINE=97386591; PubMed=9244440;
RA Rehli M., Krause S.W., Andreesen R.;
RT "Molecular characterization of the gene for human cartilage gp-39
RT (CHI3L1), a member of the chitinase protein family and marker for
RT late stages of macrophage differentiation.";
RL Genomics 43:221-225(1997).
RN [3]
RP SEQUENCE OF 22-45.
RX MEDLINE=90328983; PubMed=2475755;
RA Nyirkos P., Golds E.E.;
RI "Human synovial cells secrete a 39 kDa protein similar to a bovine
RI mammary protein expressed during the non lactating period.";
RL Biochem. J. 269:265-268(1990).
CC -!- FUNCTION: MAY PLAY AN IMPORTANT ROLE IN THE CAPACITY OF CELLS TO
CC RESPOND TO AND COPE WITH CHANGES IN THEIR ENVIRONMENT.
CC -!- SUBCELLULAR LOCATION: Extracellular.
CC -!- TISSUE SPECIFICITY: PRESENT IN ARTICULAR CHONDROCYTES, SYNOVIAL
CC CELLS AS WELL AS IN LIVER. UNDETECTABLE IN MUSCLE TISSUES, LUNG,
CC PANCREAS, MONONUCLEAR CELLS, OR FIBROBLASTS.
CC -!- PTM: GLYCOSYLATED.
CC -!- SIMILARITY: BELONGS TO FAMILY 18 OF GLYCOSYL HYDROLASES.
CC ----------
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC ----------

```
DR   EMBL; M80927; AAA16074.1;  -.
DR   EMBL; Y08374; CAA69661.1;  -.
DR   EMBL; Y08375; CAA69661.1; JOINED.
DR   EMBL; Y08376; CAA69661.1; JOINED.
DR   EMBL; Y08377; CAA69661.1; JOINED.
DR   EMBL; Y08378; CAA69661.1; JOINED.
DR   PIR; S10677; S10677.
DR   PIR; A33162; A33162.
DR   MIM; 601525; -.
DR   InterPro; IPR001579; Chitinase_2.
DR   InterPro; IPR001223; Glyco_hydro_18.
DR   Pfam; PF00704; Glyco_hydro_18; 1.
DR   PROSITE; PS01095; CHITINASE_18; FALSE_NEG.
KW   Glycoprotein; Signal.
FT   SIGNAL         1     21
FT   CHAIN         22    383       CHITINASE-3 LIKE PROTEIN 1.
FT   CARBOHYD      60     60       N-LINKED (GLCNAC...) (POTENTIAL).
SQ   SEQUENCE   383 AA;  42613 MW;  76ADD829BEEEC2D1 CRC64;


  Query Match             100.0%;  Score 43;  DB 2;  Length 383;
  Best Local Similarity   100.0%;  Pred. No. 0.045;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 FILASSEIG 9
           |||||||||
Db     265 FILASSEIG 273


RESULT   2
CHL2_HUMAN
ID   CHL2_HUMAN     STANDARD;      PRT:   390 AA.
AC   Q15782; Q15783; Q15749;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Chitinase 3-like protein 2 precursor (YKL-39) (Cartilage glycoprotein
DE   39).
GN   CHI3L2.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A. (SHORT AND LONG ISOFORMS).
RA   Grossman A., Matsuyama T., Baker E., Waterhouse P., Sutherland G.R.,
RA   Mak T.W.;
RT   "Cloning of a novel lymphoid restricted human chitinase and
RT   localization to 1p14.3.";
RL   Submitted (MAY-1996) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A. (SHORT ISOFORM).
RC   TISSUE=Articular cartilage;
RX   MEDLINE=96325055; PubMed=8702629;
RA   Hu B., Trinh K., Figueira W.F., Price P.A.;
RT   "Isolation and sequence of a novel human chondrocyte protein related
RT   to mammalian members of the chitinase protein family.";
RL   J. Biol. Chem. 271:19415-19420(1996).
CC   -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS, A SHORT FORM (BROWN BEER) AND A
CC       LONG FORM; MAY BE PRODUCED BY ALTERNATIVE SPLICING.
CC   -!- TISSUE SPECIFICITY: HIGHEST EXPRESSION IN CHONDROCYTES, FOLLOWED
CC       BY SYNOVIOCYTES, LUNG AND HEART. NOT DETECTED IN BRAIN, SPLEEN,
CC       PANCREAS, AND LIVER.
CC   -!- SIMILARITY: BELONGS TO FAMILY 18 OF GLYCOSYL HYDROLASES.
CC   -----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----------------------------------------------------------------------
DR   EMBL; U58515; AAB64534.1;  -.
DR   EMBL; U58514; AAB64533.1;  -.
DR   EMBL; U49835; AAC50597.1; ALT_INIT.
DR   MIM; 601526; -.
DR   InterPro; IPR001579; Chitinase_2.
DR   InterPro; IPR001223; Glyco_hydro_18.
DR   Pfam; PF00704; Glyco_hydro_18; 1.
DR   PROSITE; PS01095; CHITINASE_18; FALSE_NEG.
KW   Hydrolase; Signal; Alternative splicing.
FT   SIGNAL         1     26       POTENTIAL.
FT   CHAIN         27    390       CHITINASE-3 LIKE PROTEIN 2.
FT   CARBOHYD      35     35       N-LINKED (GLCNAC...) (POTENTIAL).
FT   VARSPLIC       1     12       MGAIIMDQQKSLW -> SQQLWLGSPDEVWEYYV-
FT                                 AIWDSKIREGSIDFSEHPSI (IN ISOFORM 2).
SQ   SEQUENCE   390 AA;  43500 MW;  97B86A2F3AA35077 CRC64;


  Query Match             79.1%;  Score 34;  DB 1;  Length 390;
  Best Local Similarity   87.5%;  Pred. No. 5.1;
  Matches    7;  Conservative    1;  Mismatches    0;  Indels    :  Gaps    :

Qy       1 FILASSEI 8
           |||||:||
Db     273 FILASAEI 280


RESULT   3
YACK_BACSU
ID   YACK_BACSU     STANDARD;      PRT:   360 AA.
AC   P37573;
DT   01-OCT-1994 (Rel. 30, Created)
DT   01-OCT-1994 (Rel. 30, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Hypothetical 40.7 KDa protein in MECB-GLIX intergenic region.
GN   YACK.
OS   Bacillus subtilis.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group;
OC   Bacillus/Staphylococcus group; Bacillus.
OX   NCBI_TaxID=1423;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=168;
RX   MEDLINE=96051385; PubMed=7584024;
RA   Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
RA   Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,
RA   Borriss R., Boursier L., Brans A., Braun M., Brey S.C., Buchanan J.,
RA   Bunn D.M., Castellin M., Chalker A., Chapuis J., Chen L., Chen Y.,
RA   Christiansen C., Colley A., Collins C.M., Cormier M.J., Corton C.,
RA   Crabtree J., Cutting S., Danchin A., De Biase A., Delorme C.,
RA   Dervyn E., Devine K.M., Dominguez-Cuevas P., Dubois L., Ehrlich S.D.,
RA   Emmerson P., Entian K.-D., Errington J., Fabret C., Ferrari E.,
RA   Foulger D., Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A.,
RA   Galleron N., Ghim S.-Y., Glaser P., Goffeau A., Golightly E.J.,
RA   Grandi G., Guiseppi G., Guy B.J., Haga K., Haiech J., Harwood C.R.,
RA   Henaut A., Hilbert H., Holsappel S., Hosono S., Hullo M.-F.,
RA   Itaya M., Jones L., Joris B., Karamata D., Kasahara Y.,
RA   Klaerr-Blanchard M., Klein C., Kobayashi Y., Koetter P., Koningstein G.,
RA   Krogh S., Kumano M., Kurita K., Lapidus A., Lardinois S., Lauber J.,
RA   Lazarevic V., Lee S.-M., Levine A., Liu H., Masuda S., Mauel C.,
RA   Medigue C., Medina N., Mellado R.P., Mizuno M., Moestl D.,
RA   Nakai S., Noback M., Noone D., O'Reilly M., Ogawa K., Ogiwara A.,
RA   Oudega B., Park S.-H., Parro V., Pohl T.M., Portetelle D.,
RA   Porwollik S., Prescott A.M., Presecan E., Pujic P., Purnelle B.,
RA   Rapoport G., Rey M., Reynolds S., Rieger M., Rivolta C., Rocha E.,
RA   Roche B., Rose M., Sadaie Y., Sato T., Scanlan E., Schleich S.,
RA   Schroeter R., Scoffone F., Sekiguchi J., Sekowska A., Seror S.J.,
RA   Serror P., Shin B.-S., Soldo B., Sorokin A., Tacconi E., Takagi T.,
RA   Takahashi H., Takemaru K., Takeuchi M., Tamakoshi A., Tanaka T.,
RA   Terpstra P., Tognoni A., Tosato V., Uchiyama S., Vandenbol M.,
RA   Vannier F., Vassarotti A., Viari A., Wambutt R., Wedler E.,
RA   Wedler H., Weitzenegger T., Winters P., Wipat A., Yamamoto H.,
RA   Yamane K., Yasumoto K., Yata K., Yoshida K., Yoshikawa H.-F.,
RA   Zumstein E.;
RT   "The complete genome sequence of the Gram-positive bacterium
RT   Bacillus subtilis.";
RL   Nature 390:249-256(1997).
CC   -----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----------------------------------------------------------------------
DR   EMBL; D26185; BAA05322.1;  -.
DR   EMBL; Z99104; CAB11864.1;  -.
DR   SubtiList; BG10170; yackA.
DR   InterPro; IPR003350; DUF147.
DR   InterPro; IPR000445; HHH.
DR   InterPro; IPR003583; HHH_1.
DR   Pfam; PF02457; DUF147; 1.
DR   Pfam; PF00633; HHH; 1.
DR   SMART; SM00278; HhH1; 1.
KW   Hypothetical protein; Complete proteome.
SQ   SEQUENCE   363 AA;  40754 MW;  [illegible] CRC64;


  Query Match             74.4%;  Score 32;  DB 1;  Length 360;
  Best Local Similarity   75.0%;  Pred. No. 17;
  Matches    6;  Conservative    2;  Mismatches    0;  Indels    :  Gaps    :
```

```
Qy      2 FTLASSETG 9
          : || ::|||
Db     70 FNLAEAETG 87

RESULT   4
CHI1_MOUSE
ID   CHI1_MOUSE     STANDARD;      PRT:   381 AA.
AC   P36366;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Chitinase 3-like protein 1 precursor (Cartilage glycoprotein-39)
DE   (GP-39) (BRP-39 protein).
GN   CHI3L1 OR BRP39.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=FVB/N; TISSUE=Breast;
RX   MEDLINE=95060797; PubMed=7970700;
RA   Morrison B.W., Cooper D.;
RT   "hcm and ras initiate murine mammary tumors that share genetic markers
RT   generally absent in c-myc and int-2 initiated tumors.";
RL   Oncogene 9:3417-3424(1994).
CC   -!- FUNCTION: MAY PLAY AN IMPORTANT ROLE IN THE CAPACITY OF CELLS TO
CC       RESPOND TO AND COPE WITH CHANGES IN THEIR ENVIRONMENT (BY
CC       SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: Extracellular (By similarity).
CC   -!- SIMILARITY: BELONGS TO FAMILY 18 OF GLYCOSYL HYDROLASES.
CC   ---------- ---- ---------- -- ---------- ---- ---- ---------- ----------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------- ---- ---------- -- ---------- ---- ---- ---------- ----------
DR   EMBL; X66085; CAA46744.1; -.
DR   MGD; MGI:1340??; Chi1l1.
DR   InterPro; IPR001579; Chitinase_2.
DR   InterPro; IPR001223; Glyco_hydro_18.
DR   Pfam; PF00704; Glyco_hydro_18; 1.
DR   PROSITE; PS01095; CHITINASE_18; FALSE_NES.
KW   Glycoprotein; Signal.
FT   SIGNAL      1     21       POTENTIAL.
FT   CHAIN      22    381       CHITINASE 3-LIKE PROTEIN 1.
FT   CARBOHYD   ??    ??        N-LINKED (GLCNAC...) (POTENTIAL).
SQ   SEQUENCE   381 AA.  43??1 MW.  E?C?1?B1641?=4?3 CRC64;


  Query Match             74.4%;  Score 42;  DB 1;  Length 381;
  Best Local Similarity   100.0%;  Pred. No. 14;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 FTLASSF 7
          : || |||
Db    226 FTLASSF 232


RESULT   5
Y034_BPHI1
ID   Y034_BPHI1     STANDARD;      PRT:   187 AA.
AC   P51738;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   15-OCT-2001 (Rel. 40, Last annotation update)
DE   Hypothetical 21.8 kDa protein in LYS 3'region (ORF34).
OS   Bacteriophage BPI.
OC   Viruses; dsDNA viruses, no RNA stage; Caudovirales; Myoviridae.
OX   NCBI_TaxID=10690;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BPIC1;
RX   MEDLINE=96279738; PubMed=8710508;
RA   Esposito D., Fitzmaurice W.P., Benjamin R.C., Goodman S.D.,
RA   Waldman A.S., Scocca J.J.;
RT   "The complete nucleotide sequence of bacteriophage BPI DNA.";
RL   Nucleic Acids Res. 24:2360-2368(1996).
CC   ---------- ---- ---------- -- ---------- ---- ---- ---------- ----------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------- ---- ---------- -- ---------- ---- ---- ---------- ----------
DR   EMBL; U24159; AAB09221.1; -.
KW   Hypothetical protein.
SQ   SEQUENCE   187 AA;  20?0? MW;  A?A?C?G?TDEC?E? CRC64;


  Query Match             72.1%;  Score 41;  DB 1;  Length 187;
  Best Local Similarity   66.7%;  Pred. No. 10;
  Matches    6;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 FTLASSETG 9
          : || ::|||
Db     72 FNLAEAETG 80


RESULT   6
INHA_MYCAV
ID   INHA_MYCAV     STANDARD;      PRT:   268 AA.
AC   O07400;
DT   15-DEC-1998 (Rel. 37, Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   15-DEC-2000 (Rel. 37, Last annotation update)
DE   Enoyl [acyl carrier protein] reductase [NADH] (EC 1.3.1.9) (NADH
DE   dependent enoyl-ACP reductase).
GN   INHA.
OS   Mycobacterium avium.
OC   Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC   Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX   NCBI_TaxID=1764;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=GIR10;
RX   MEDLINE=98195279; PubMed=9534249;
RA   Lalo M., Gusberti L., de Kossai E., Speziale L., Riccardi G.;
RT   "Determination of a 1545 bp nucleotide sequence around the InhA gene
RT   of Mycobacterium avium and similarity analysis of the products of
RT   putative ORFs.";
RL   Microbiology 144:807-814(1998).
CC   -!- CATALYTIC ACTIVITY: Acyl [acyl carrier protein] + NAD(+) = trans-
CC       2,3-dehydroacyl-[acyl-carrier protein] + NADH.
CC   -!- PATHWAY: SECOND REDUCTIVE STEP IN FATTY ACID BIOSYNTHESIS. THIS
CC       ENZYME IS INVOLVED IN MYCOLIC ACID BIOSYNTHESIS.
CC   -!- SUBUNIT: HOMOTETRAMER (BY SIMILARITY).
CC   -!- SIMILARITY: SOME SIMILARITY TO THE SHORT CHAIN
CC       DEHYDROGENASES/REDUCTASES (SDR) FAMILY.
CC   ---------- ---- ---------- -- ---------- ---- ---- ---------- ----------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------- ---- ---------- -- ---------- ---- ---- ---------- ----------
DR   EMBL; AF002134; AAC46204.1; -.
DR   HSSP; P45533; 1ENY.
```

```
DR   InterPro; IPR002198; ADH_short.
DR   Pfam; PF00106; adh_short; 1.
KW   Oxidoreductase; NAD; Fatty acid biosynthesis.
FT   NP_BIND     145     164       NAD (POTENTIAL).
SQ   SEQUENCE   260 AA;  28529 MW;  F73501BD2B0F9990 CRC64;


  Query Match             72.1%;  Score 41;  DB 1;  Length 260;
  Best Local Similarity   75.0%;  Pred. No. 15;
  Matches    6;  Conservative    2;  Mismatches    0;  Indels    1;  Gaps    0;

Qy       2 TLASSEIG 9
             ::|.|.|:
Db      94 SIASSEIG 101


RESULT   7
CN7B_SCHPO
ID   CN7B_SCHPO     STANDARD;      PRT:   142 AA.
AC   Q09722; Q9US29;
DT   01-NOV-1995 (Rel. 32, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Cop9/signalosome subunit 7B.
GN   CSN7B OR SPAC01A2.01 OR SPAC1751.03.
OS   Schizosaccharomyces pombe (Fission yeast).
OC   Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC   Schizosaccharomycetales; Schizosaccharomycetaceae;
OC   Schizosaccharomyces.
OX   NCBI_TaxID=4896;
RN   [1]
RP   SEQUENCE OF 1-378 FROM N.A.
RC   STRAIN=972;
RA   Zimmerman W., Wambutt R., McDonnell R.C., Rajandream M.A.;
RA   Barrell B.G.;
RL   Submitted (NOV-1999) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE OF 320-402 FROM N.A.
RC   STRAIN=972;
RA   Devlin K., Chambers C.M., Barrell B.G., Rajandream M.A., Wilson V.;
RL   Submitted (JUL-1995) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: CONTAINS 1 PCI DOMAIN.
CC   -----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----------------------------------------------------------------------
DR   EMBL; AL133154; CAB61449.1; -.
DR   EMBL; Z50213; CAA90459.1; -.
DR   InterPro; IPR000717; PCI.
DR   SMART; SM00088; PINT; 1.
KW   Signalosome.
FT   DOMAIN      274    365       PCI.
SQ   SEQUENCE   402 AA;  45085 MW;  CD972E781B6F90A8 CRC64;


  Query Match             72.1%;  Score 41;  DB 1;  Length 402;
  Best Local Similarity   66.7%;  Pred. No. 25;
  Matches    6;  Conservative    2;  Mismatches    1;  Indels    1;  Gaps    0;

Qy       1 FIASSEIG 9
             |:|:||:|
Db     383 FIASSKKG 391


RESULT   8
FLGL_VIBPA
ID   FLGL_VIBPA     STANDARD;      PRT:   445 AA.
AC   Q03475;
DT   01-OCT-1993 (Rel. 27, Created)
DT   01-OCT-1993 (Rel. 27, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Lateral flagellar hook associated protein 2 (HAP2) (Filament 'A-
DE   protein) (Flagellar cap protein).
GN   FLGL OR LAFB.
OS   Vibrio parahaemolyticus.
OC   Bacteria; Proteobacteria; Gamma subdivision; Vibrionaceae; Vibrio.
OX   NCBI_TaxID=670;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BB22;
RX   MEDLINE=93274702; PubMed=8501040;
RA   McCarter L.L., Wright M.E.;
RT   "Identification of genes encoding components of the swarmer cell
RT   flagellar motor and propeller and a sigma factor controlling
RT   differentiation of Vibrio parahaemolyticus.";
RL   J. Bacteriol. 175:3361-3371(1993).
CC   -!- FUNCTION: REQUIRED FOR THE MORPHOGENESIS AND ASSEMBLY OF THE
CC       THE FLAGELLAR FILAMENT BY FACILITATING POLYMERIZATION OF
CC       FLAGELLIN MONOMERS AT THE TIP OF GROWING FILAMENT. FORMS A DEFINED
CC       STRUCTURE, WHICH PREVENTS FLAGELLIN SUBUNITS (TRANSIT ACROSS THE
CC       THE CENTRAL CHANNEL OF THE FLAGELLUM) FROM LEAKING OUT INTO THE
CC       EXTRACELLULAR MEDIUM AT THE GROWING END. ESSENTIAL FOR SWARMING M
CC   -!- SUBUNIT: HOMOPENTAMER (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: FLAGELLAR.
CC   -!- MISCELLANEOUS: VIBRIOPHAGE MYOVIRUS POSSESSES TWO FLAGELLAR T-TYPE
CC       A SINGLE POLAR FLAGELLUM PROPELS THE BACTERIUM IN LIQUID
CC       (SWIMMING), WHILE MULTIPLE LATERAL FLAGELLA (FLAGELLIN DIFFERENT
CC       THE BACTERIUM OVER SURFACES (SWARMING). THE POLAR FLAGELLUM IS
CC       SYNTHESIZED CONSTITUTIVELY BUT LATERAL FLAGELLA ARE EXPRESSED ONLY
CC       UNDER CONDITIONS IN WHICH THE POLAR FLAGELLUM IS NOT FUNCTIONAL.
CC   -!- SIMILARITY: BELONGS TO THE FLIC FAMILY.
CC   -----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----------------------------------------------------------------------
DR   EMBL; L06176; AAA27540.1; -.
DR   EMBL; Y52957; AAB07451.1; -.
DR   PIR; B40590; B40590.
DR   InterPro; IPR001491; FliD.
DR   Pfam; PF02465; FliD; 1.
KW   Flagellar; Coiled coil.
FT   COILED      [illegible]
SQ   SEQUENCE   445 AA;  48264 MW;  0E7F7089A5258052A CRC64;


  Query Match             72.1%;  Score 41;  DB 1;  Length 445;
  Best Local Similarity   66.7%;  Pred. No. 28;
  Matches    6;  Conservative    2;  Mismatches    1;  Indels    1;  Gaps    0;

Qy       1 FIASSEIG 9
             |:|:||:|
Db     177 FMLSSIEG 185


RESULT   9
VL2_HPV6A
ID   VL2_HPV6A      STANDARD;      PRT:   499 AA.
AC   Q84297;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   15-JUL-1998 (Rel. 36, Last annotation update)
DE   Minor capsid protein L2.
GN   L2.
OS   Human papillomavirus type 6a.
OC   Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC   Papillomavirus.
```

```
RX   NCBI_TaxID=37122;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=96120192; PubMed=7778243;
RA   Heilmann K.L., Cook J.C., Coyne J.C., Brown L.K., Schmitz L.D.,
RA   George H.A., Rosolowsky M., Fife K.H., Jansen K.U.;
RT   "Sequence determination of human papillomavirus type 58 and assembly
RT   of virus like particles in Saccharomyces cerevisiae.";
RL   Virology 205:506-518(1994).
CC   -!- ------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ------------------------------------------------------------------
DR   EMBL; L41216; AAA43217.1; -.
DR   InterPro; IPR000984; late_L2.
DR   Pfam; PF00518; late_protein_L2; 1.
KW   Coat protein; late protein.
SQ   SEQUENCE   459 AA;  49406 MW;  FA1B2C7ACBC89406 CRC64;


  Query Match             72.1%;  Score 41;  DB 1;  Length 459;
  Best Local Similarity   75.0%;  Pred. No. 29;
  Matches    6;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 FILASSET 8
           |: ||||
Db     129 FTIISSET 136


RESULT  3
VL2_BPV9B
ID   VL2_BPV9B      STANDARD;      PRT:   459 AA.
AC   P04116;
DT   21-JUL-1986 (Rel. 1, Created)
DT   21-JUL-1986 (Rel. 1, Last sequence update)
DT   01-OCT-1996 (Rel. 34, Last annotation update)
DE   Minor capsid protein L2.
GN   L2.
OS   Human papillomavirus type 9b.
OC   Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC   Papillomavirus.
OX   NCBI_TaxID=10649;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=84142940; PubMed=6321572;
RA   Schwarz E., Durst M., Demankowski C., Lattermann G., Zech R.,
RA   Wolfsperger E., Suhai S., zur Hausen H.;
RT   "DNA sequence and genome organization of genital human papillomavirus
RT   type 6b.";
RL   EMBO J. 2:2341-2348(1983).
CC   -!- ------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ------------------------------------------------------------------
DR   EMBL; X00203; CAA25025.1; -.
DR   PIR; A03647; VLWS6.
DR   InterPro; IPR000984; late_L2.
DR   Pfam; PF00518; late_protein_L2; 1.
KW   Coat protein; late protein.
SQ   SEQUENCE   459 AA;  49491 MW;  15048EB890C48E45 CRC64;


  Query Match             72.1%;  Score 41;  DB 1;  Length 459;
  Best Local Similarity   75.0%;  Pred. No. 29;
  Matches    6;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 FILASSET 8
           |: ||||
Db     129 FTIISSET 136


RESULT  11
VG29_BPT4
ID   VG29_BPT4      STANDARD;      PRT:   590 AA.
AC   P13337;
DT   01-JAN-1990 (Rel. 13, Created)
DT   01-JAN-1990 (Rel. 13, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
DE   Tail tube assembly protein gp29 (tail length regulator).
DE   (Folylpolyglutamate synthase (EC 6.3.2.17)).
GN   29.
OS   Bacteriophage T4.
OC   Viruses; dsDNA viruses, no RNA stage; Caudovirales; Myoviridae;
OC   T4-like phages.
OX   NCBI_TaxID=10665;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=88206086; PubMed=3363870;
RA   Ishimoto L.K., Ishimoto K.S., Cascino A., Cipollaro M.,
RA   Eiserling F.A.;
RT   "The structure of three bacteriophage T4 genes required for tail-tube
RT   assembly.";
RL   Virology 164:81-90(1988).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Kutter E., Ariaka F., Kamisawa T., Ismita A., Mosig G.,
RA   Mosyanzhinov V., Ruger W., Stidham T., Thomas E.;
RT   "Bacteriophage T4 genome analysis.";
RL   Submitted (JUL-2000) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: STRUCTURAL COMPONENT OF THE BASEPLATE OF THE PHAGE AND
CC       FUNCTIONS AS A STRUCTURAL INITIATOR OF THE CENTRAL HUB ASSEMBLY OF
CC       THE BASEPLATE. IT ALSO CATALYZES FOLYL POLYGLUTAMATE SYNTHETASE
CC       ACTIVITY.
CC   -!- CATALYTIC ACTIVITY: ATP + TETRAHYDROFOLYL-[GLU][N] + L-GLUTAMATE =
CC       ADP + ORTHOPHOSPHATE + TETRAHYDROFOLYL-[GLU][N+1].
CC   -!- ------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ------------------------------------------------------------------
DR   EMBL; M20298; AAA42538.1; -.
DR   EMBL; AF158101; AAD42447.1; -.
DR   PIR; JF0035; GBBPT4.
KW   Structural protein; Late protein; Ligase.
SQ   SEQUENCE   590 AA;  64410 MW;  FBE6161019E6FBB4 CRC64;


  Query Match             72.1%;  Score 41;  DB 1;  Length 590;
  Best Local Similarity   77.8%;  Pred. No. 39;
  Matches    7;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 FILASSETG 9
           || |||||
Db     278 FTKFSSETG 286


RESULT  12
POL1_CHICK
ID   POL1_CHICK     STANDARD;      PRT:   246 AA.
AC   P10399;
DT   01-MAR-1989 (Rel. 10, Created)
DT   01-MAR-1989 (Rel. 10, Last sequence update)
```

```
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Retrovirus-related Pol polyprotein [Contains: Reverse transcriptase
DE   (EC 2.7.7.49); Endonuclease] (Fragments).
GN   Pol.
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=9031;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE 85,264471; PubMed=2410912;
RA   Dunwiddie C., Faras A.J.;
RT   "Presence of retrovirus reverse transcriptase related gene sequences
RT   in avian cells lacking endogenous avian leukosis viruses.";
RL   Proc. Natl. Acad. Sci. U.S.A. 82:5097-5101(1985).
RN   [2]
RP   REVISION TO 2.
RA   Dunwiddie C., Faras A.J.;
RL   Submitted (OCT 1985) to the EMBL/GenBank/DDBJ databases.
CC   -!- ---------- ---- ---- ---- ---- ---- ---- ---- ----
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------- ---- ---- ---- ---- ---- ---- ---- ----
DR   EMBL; K03271; -; NOT_ANNOTATED_CDS.
DR   EMBL; K03272; -; NOT_ANNOTATED_CDS.
DR   PIR; A25779; A25779.
DR   Pfam; PF00665; rve; 1.
KW   Hydrolase; Transferase; RNA-directed DNA polymerase; Nuclease;
KW   Endonuclease; Polyprotein.
FT   NON_TER       1       1
FT   NON_CONS    111     112
FT   NON_TER     246     246
SQ   SEQUENCE   246 AA;  26854 MW;  EA036F7C07920706 CRC64;


  Query Match             69.8%;  Score 30;  DB 1;  Length 246;
  Best Local Similarity   77.8%;  Pred. No. 24;
  Matches    7;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 FTLASSFIG 9
           | :| :|
Db      98 FTDASSSIG 106


RESULT  13
MSHR_BOVIN
ID   MSHR_BOVIN     STANDARD;      PRT:   317 AA.
AC   P47798; Q28025;
DT   01-FEB-1995 (Rel. 31, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
DE   Melanocyte-stimulating hormone receptor (MSH-R) (Melanotropin
DE   receptor) (Melanocortin 1 receptor) (MC1-R) (MCR).
GN   MC1R OR MSHR.
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC   Bovidae; Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE 94307498; PubMed=8034852;
RA   Vanetti M., Schoonmaker C., Meyerhof W., Hoellt V.;
RT   "Molecular cloning of a bovine MSH receptor which is highly expressed
RT   in the testes.";
RL   FEBS Lett. 348:268-272(1994).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Holstein;
RX   MEDLINE 96325106; PubMed=8661736;
RA   Joerg H., Fries H.R., Meijerink E., Stranzinger G.F.;
RT   "Red coat color in Holstein cattle is associated with a polymorphism of
RT   the MSHR gene.";
RL   Mamm. Genome 7:317-318(1996).
RN   [3]
RP   SEQUENCE FROM N.A.
RA   Klungland H., Rood K.H., Vage D.I.;
RL   Submitted (OCT 1997) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Holstein;
RA   Kriegesmann B., Brenig B., Bierkes B., Jansen S.;
RT   "A new melanocortin 1 receptor allele is coupled with Agouti and
RT   colour pattern in Holstein cattle.";
RL   Submitted (JUN 1999) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: RECEPTOR FOR MSH (ALPHA, BETA) AND ACTH. ACTIVE N [illegible]
CC       BY ACTIVE WITH SOMEWHAT MORE THE ACTIVITY OF THIS RECEPTOR IS
CC       MEDIATED BY G PROTEINS WHICH ACTIVATE ADENYLATE CYCLASE. MAY BE
CC       INVOLVED IN SPERMATOGENESIS.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN THE TESTIS.
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G PROTEIN-COUPLED RECEPTORS.
CC   ---------- ---- ---- ---- ---- ---- ---- ---- ----
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------- ---- ---- ---- ---- ---- ---- ---- ----
DR   EMBL; S71017; AAB31361.1; -.
DR   EMBL; U39469; AAC48590.2; -.
DR   EMBL; Y13957; CAA74291.1; -.
DR   EMBL; Y19104; CAB64418.1; -.
DR   GCRDb; GCR_1220; -.
DR   GCRDb; GCR_1237; -.
DR   GCRDb; GCR_2481; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS00262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein;
KW   Phosphorylation; Lipoprotein; Palmitate.
FT   DOMAIN        1     37       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     38     57       1 (POTENTIAL).
FT   DOMAIN       64     72       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     73     93       2 (POTENTIAL).
FT   DOMAIN       94    118       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    119    140       3 (POTENTIAL).
FT   DOMAIN      141    163       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    164    183       4 (POTENTIAL).
FT   DOMAIN      184    191       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    192    211       5 (POTENTIAL).
FT   DOMAIN      212    243       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    241    266       6 (POTENTIAL).
FT   DOMAIN      267    279       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    280    300       7 (POTENTIAL).
FT   DOMAIN      301    317       CYTOPLASMIC (POTENTIAL).
FT   CARBOHYD     29     29       N-LINKED (GLCNAC...) (POTENTIAL).
FT   LIPID       315    315       PALMITATE (POTENTIAL).
FT   CONFLICT     40     41       IP -> SI (IN REF. 1).
FT   CONFLICT     99     99       S -> R (IN REF. 1 AND 2).
SQ   SEQUENCE   317 AA;  34916 MW;  1C39926141F7EF10 CRC64;


  Query Match             69.8%;  Score 30;  DB 1;  Length 317;
  Best Local Similarity   66.7%;  Pred. No. 32;
  Matches    6;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;
```

```
Qy       1 FTLASSETC 9
             |||| : ||
Db     24 FTLAINKET 2

RESULT  14
SM4B_MOUSE
ID   SM4B_MOUSE      STANDARD;      PRT:   782 AA.
AC   Q62175;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
DE   Semaphorin 4B (Semaphorin C) (Sema C) (Fragment).
GN   SEMA4B OR SEMAC OR SEMC.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=NMRI; TISSUE=Brain;
RX   MEDLINE=96267433; PubMed=7748561;
RA   Puschel A.W., Adams R.H., Betz H.;
RT   "Murine semaphorin D/collapsin is a member of a diverse gene family
RT   and creates domains inhibitory for axonal extension.";
RL   Neuron 14:941-948(1995).
RN   [2]
RP   INTERACTION WITH GIPC.
RX   MEDLINE=99276778; PubMed=10330131;
RA   Wang L.H., Katz R.S., Strittmatter S.M.;
RT   "A PDZ protein regulates the distribution of the transmembrane
RT   semaphorin, M-SemF.";
RL   J. Biol. Chem. 274:14137-14140(1999).
CC   -!- FUNCTION: INHIBITS AXONAL EXTENSION BY PREVENTING LAMELLIPODIA TO
CC       SPREAD BEHIND SITES INACCESSIBLE TO GROWING AXONS.
CC   -!- SUBUNIT: INTERACTS WITH GIPC PDZ DOMAIN.
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC   -!- DEVELOPMENTAL STAGE: EXPRESSED FROM DAY 10 IN THE BRAIN. LOW
CC       LEVELS FOUND BETWEEN DAYS 10 TO 12. EXPRESSION PEAKS ON DAY 15 WITH
CC       MODERATE LEVELS FROM THEN UNTIL BIRTH.
CC   -!- SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.
CC   -!- SIMILARITY: CONTAINS 1 SEMA DOMAIN.
CC   -!- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAIN.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; X85992; CAA59884.1; -.
DR   MGD; MGI:107559; Sema4b.
DR   InterPro; IPR001627; PSI.
DR   InterPro; IPR002165; Plexin repeat.
DR   InterPro; IPR001627; Sema.
DR   Pfam; PF01437; Plexin repeat; 1.
DR   Pfam; PF01403; Sema; 1.
DR   SMART; SM00423; PS; 1.
KW   Transmembrane; Immunoglobulin domain; Multigene family; Neurogenesis;
KW   Developmental protein; Glycoprotein.
FT   NON_TER       1      1
FT   DOMAIN        1    562       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    663    683       POTENTIAL.
FT   DOMAIN      684    782       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN       20    492       SEMA.
FT   DOMAIN      543    608       IG-LIKE C2-TYPE DOMAIN.
FT   DOMAIN      703    726       PRO RICH.
FT   DISULFID     58    601       BY SIMILARITY.
FT   CARBOHYD     12     12       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     19     19       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     42     42       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     88     88       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    356    356       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    471    471       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    526    526       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    574    574       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    639    639       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   782 AA;  86823 MW;  627A81EC8EF8F7AC8 CRC64;


  Query Match             69.8%;  Score 30;  DB 1;  Length 782;
  Best Local Similarity   66.7%;  Pred. No. 92;
  Matches    6;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 FTLASSETC 9
             |||| : ||
Db     112 FTLAQDEAG 120


RESULT  15
STBB_PSESM
ID   STBB_PSESM      STANDARD;      PRT:   139 AA.
AC   Q52562;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   01-NOV-1997 (Rel. 35, Last annotation update)
DE   Plasmid stability protein stbB.
GN   STBB.
OS   Pseudomonas syringae (pv. tomato).
OC   bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC   Pseudomonas.
OX   NCBI_TaxID=323;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Hackkamp T., Kobayashi K., Hayes C., Staskawicz M.M.;
RL   Submitted (JUL-1997) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: INVOLVED IN PLASMID STABILITY.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; L48985; AAB81645.1; -.
DR   InterPro; IPR002716; PIN.
DR   Pfam; PF01850; PIN; 1.
SQ   SEQUENCE   139 AA;  15156 MW;  9F4281705D362D67 CRC64;


  Query Match             67.4%;  Score 29;  DB 1;  Length 139;
  Best Local Similarity   55.6%;  Pred. No. 21;
  Matches    5;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 FTLASSEIG 9
             |:|: :||
Db     116 FTVATRDIG 124


Search completed: August 26, 2002, 06:29:02
Job time: 928 sec
```

Geneticc version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

```
SM protein - protein search, using sw model

Run on:          August 26, 2002, 06:28:04 ; Search time 13.24 Seconds
                                            (without alignments)
                                            11.959 Million cell updates/sec

Title:           US-09-744-282-2
Perfect score:   441
Sequence:        1 FILASSEIG 9

Scoring table:   BLOSUM62
                 Gapop 10.0 , Gapext 0.5

Searched:        562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters:      562222

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       SPTREMBL_19:*
                 1: sp_archea:*
                 2: sp_bacteria:*
                 3: sp_fungi:*
                 4: sp_human:*
                 5: sp_invertebrate:*
                 6: sp_mammal:*
                 7: sp_mhc:*
                 8: sp_organelle:*
                 9: sp_phage:*
                 10: sp_plant:*
                 11: sp_rodent:*
                 12: sp_virus:*
                 13: sp_vertebrate:*
                 14: sp_molass_fish:*
                 15: sp_virus:*
                 16: sp_bacteriap:*
                 17: sp_archaep:*

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.
```

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	441	100.0	383	4	Q96H17	Q96H17 homo sapien
2	87	86.0	441	2	O52483	O52483 serratia ma
3	85	81.4	481	2	Q56I9	Q56I9 pseudomonas
4	84	79.2	390	4	Q96E97	Q96E97 homo sapien
5	83	79.7	393	7	Q29411	Q29411 sus scrofa
6	83	76.7	1265	4	Q9P2Q7	Q9P2Q7 homo sapien
7	82	74.4	251	17	O27403	O27403 methanother
8	82	74.4	261	17	Q9HPA0	Q9HPA0 halobacteri
9	82	74.4	352	11	Q9WTV1	Q9WTV1 rattus norv
10	82	74.4	372	13	Q98Y58	Q98Y58 brachydanio
11	82	74.4	381	12	Q99J84	Q99J84 mus musculu
12	82	74.4	446	10	Q9SXN2	Q9SXN2 pyrus pyrif
13	82	74.4	574	16	Q9SUQ2	Q9SUQ2 arabidopsis
14	82	74.4	1243	5	Q9N6T8	Q9N6T8 leishmania
15	81	72.1	186	9	Q94MX7	Q94MX7 haemophilus
16	81	72.1	202	5	Q02824	Q02824 taetypleura
17	81	72.1	202	16	Q9A9W9	Q9A9W9 caul bacter
18	81	72.1	314	4	Q9UDQ5	Q9UDQ5 homo sapien
19	81	72.1	318	5	Q18800	Q18800 caenorhabdi
20	81	72.1	324	5	O45339	O45339 caenorhabdi
21	81	72.1	373	2	Q93SR2	Q93SR2 pseudomonas
22	81	72.1	373	16	Q910X8	Q910X8 pseudomonas
23	81	72.1	414	3	Q9URK6	Q9URK6 schizosacch
24	81	72.1	420	5	Q22485	Q22485 caenorhabdi
25	81	72.1	459	12	Q99A90	Q99A90 human papil
26	81	72.1	459	12	Q99A89	Q99A89 human papil
27	81	72.1	459	12	Q9QLN9	Q9QLN9 human papil
28	81	72.1	491	5	Q9X0F6	Q9X0F6 caenorhabdi
29	81	72.1	515	2	Q9ZBD4	Q9ZBD4 rhodococcus
30	81	72.1	605	17	O26511	O26511 methanother
31	81	72.1	637	10	Q9M7C9	Q9M7C9 abies grand
32	81	72.1	817	12	Q9J5H1	Q9J5H1 fowlpox vir
33	81	72.1	817	12	Q90761	Q90761 fowlpox vir
34	81	72.1	860	10	Q9LXU3	Q9LXU3 arabidopsis
35	81	72.1	861	10	Q9LXU4	Q9LXU4 arabidopsis
36	81	72.1	1073	2	Q91SV8	Q91SV8 prevotella
37	81	72.1	1112	11	Q9WVQ1	Q9WVQ1 mus musculu
38	81	72.1	1114	11	Q9R271	Q9R271 rattus norv
39	81	72.1	1277	11	O88382	O88382 rattus norv
40	81	72.1	1278	4	O60434	O60434 homo sapien
41	81	72.1	1455	4	O60510	O60510 homo sapien
42	81	72.1	4809	2	Q93BH0	Q93BH0 streptomyce
43	80	69.8	155	6	Q95KW1	Q95KW1 bos taurus
44	80	69.8	241	3	Q12580	Q12580 chaetomium
45	80	69.8	268	16	O69626	O69626 syn bacteri

ALIGNMENTS

```
RESULT   1
Q96H17
ID   Q96H17      PRELIMINARY;      PRT:   383 AA.
AC   Q96H17;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   SIMILAR TO CHITINASE 3 LIKE 1 (CARTILAGE GLYCOPROTEIN-39).
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=BRAIN, AND GLIOBLASTOMA WITH EGFR AMPLIFICATION;
RA   Strausberg R.;
RL   Submitted (MAY-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC008568; AAH08568.1; -.
SQ   SEQUENCE   383 AA;  42625 MW;  6C9EF1338B8C7C2II CRC64;
```

```
  Query Match          100.0%;  Score 441;  DB 4;  Length 383;
  Best Local Similarity  100.0%;  Pred. No. 0.37;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 FILASSEIG 9
           |||||||||
Db     265 FILASSEIG 273


RESULT   2
O52483
ID   O52483      PRELIMINARY;      PRT:   441 AA.
AC   O52483;
DT   01-JUN-1998 (TrEMBLrel. 06, Created)
DT   01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   PUTATIVE ATP-BINDING PROTEIN.
GN   WZI.
```

OS Serratia marcescens.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Serratia.
OX NCBI_TaxID=615;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN N.ATCC;
RX MEDLINE=99172013; PubMed=10074083;
RA Sakai K., Climent N., Pique N., Sanchez C., Merino S., Nobles X.,
RA Aguilar A., Tomas J.M., Regue M.;
RT "Gene'ic analysis of the Serratia marcescens N28b O4 antigen gene
RT cluster.";
RL J. Bacteriol. 181:1883-1891(1999).
DR EMBL; AF039810; AAC00182.1; -.
DR InterPro; IPR003593; AAA.
DR InterPro; IPR003439; ABC_transporter.
DR InterPro; IPR001687; ATP_GTP_A.
DR Pfam; PF00005; ABC_tran; 1.
DR SMART; SM00382; AAA; 1.
KW ATP-binding.
SQ SEQUENCE 441 AA; 48197 MW; 56F7B3F1BDC2AB5D CRC64;

Query Match 86.0%; Score 37; DB 2; Length 441;
Best Local Similarity 100.0%; Pred. No. 9;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FILASSET 8
||||||||
Db 432 FILASSET 439

RESULT 3
O05619
ID O05619 PRELIMINARY; PRT; 481 AA.
AC O05619;
DT 01-JUL-1997 (TrEMBLrel. 04, Created)
DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE VANILLIN DEHYDROGENASE.
GN VDH.
OS Pseudomonas sp.
OC Bacteria; Proteobacteria.
OX NCBI_TaxID=306;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN HR199 (DSM 7063);
RX MEDLINE=97252489; PubMed=9098058;
RA Priefert H., Rabenhorst J., Steinbuechel A.;
RT "Molecular characterization of genes of Pseudomonas sp. strain HR199
RT involved in bioconversion of vanillin to protocatechuate.";
RL J. Bacteriol. 179:2595-2607(1997).
DR EMBL; Y11520; CAA72286.1; -.
DR HSSP; P51977; 1BXS.
DR InterPro; IPR012080; Aldehyde_dehydr.
DR Pfam; PF00171; aldedh; 1.
DR PROSITE; PS00687; ALDEHYDE_DEHYDR_GLU; UNKNOWN; 1.
SQ SEQUENCE 481 AA; 50836 MW; 966EEF29F0F0D59E CRC64;

Query Match 81.4%; Score 35; DB 2; Length 481;
Best Local Similarity 77.8%; Pred. No. 27;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FILASSET 8
|| ||||||
Db 83 FIAAASEIG 91

RESULT 4
Q96F07
ID Q96F07 PRELIMINARY; PRT; 390 AA.
AC Q96F07;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE CHITINASE 3-LIKE 2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BRAIN, AND OLF BLAST MA WITH ESTP AMPLIFIABLE N;
RA Strausberg R.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DDBJ databases.
DR EMBL; BC011460; AAH11460.1; -.
SQ SEQUENCE 390 AA; 43529 MW; 59C1E1AF,1AD49F CRC64;

Query Match 79.1%; Score 34; DB 4; Length 390;
Best Local Similarity 87.5%; Pred. No. 46;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FILASSET 8
|||||:|
Db 273 FILASAET 280

RESULT 5
Q29411
ID Q29411 PRELIMINARY; PRT; 383 AA.
AC Q29411;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE 38 KDA HEPARIN BINDING GLYCOPROTEIN.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=SMOOTH MUSCLE;
RX MEDLINE=95286589; PubMed=7769092;
RA Shackelton L.M., Mann D.M., Millis A.J.;
RT "Identification of a 38 kDa heparin binding glycoprotein (gp38k) in
RT differentiating vascular smooth muscle cells as a member of a group
RT proteins associated with tissue remodeling.";
RL J. Biol. Chem. 270:13076-13083(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=SMOOTH MUSCLE;
RA Shackelton L.M., Mann D.M., Millis A.J.;
RL Submitted (JAN-1995) to the EMBL/GenBank/DDBJ databases.
DR EMBL; U19900; AAA86482.1; -.
DR EMBL; Z47803; CAA87764.1; -.
DR InterPro; IPR001223; Glyco_hydro_18; 1.
DR Pfam; PF00704; Glyco_hydro_18; 1.
KW Heparin binding.
SQ SEQUENCE 383 AA; 42443 MW; 4D1C99B499190BC4 CRC64;

Query Match 76.2%; Score 33; DB 4; Length 383;
Best Local Similarity 87.5%; Pred. No. 59;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FILASSET 8
|||||||:
Db 265 FILASSKT 272

RESULT 6
Q9P2G7
ID Q9P2G7 PRELIMINARY; PRT; 1265 AA.
AC Q9P2G7;

ID [illegible] 200- (TrEMBLrel. 15, Created)
DT [illegible] 200- (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2001 (TrEMBLrel. 19, Last annotation update)
DE KIAA1980 PROTEIN (FRAGMENT).
GN KIAA1980.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BRAIN.
RX MEDLINE=20181126; PubMed=10716198;
RA Nagase T., Kikuno R., Ishikawa K.-I., Hirosawa M., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes.XVI.
RT The complete sequences of 150 new cDNA clones from brain which code
RT for large proteins in vitro.";
RL DNA Res. 7:65-73(2000).
DR EMBL; AB037838; BAA95618.1; -.
DR InterPro; IPR000547; mjc.
DR Pfam; PF02374; ja; 1.
FT NON_TER 1 1
SQ SEQUENCE 1165 AA; 131291 MW; [illegible] CRC64;

Query Match 96.7%; Score 14; DB 4; Length 1265;
Best Local Similarity 87.5%; Pred. No. 2.1e+02;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 FIASSEIG 9
|:|||:|||
Db 184 IIASSKFG 196

RESULT 7
Q27403
ID Q27403 PRELIMINARY; PRT; 251 AA.
AC Q27403;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE CONSERVED PROTEIN
GN MTH1449.
OS Methanothermobacter thermautotrophicus.
OC Archaea; Euryarchaeota; Methanobacteriales; Methanobacteriaceae;
OC Methanothermobacter.
OX NCBI_TaxID=145262;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN DELTA H;
RX MEDLINE=98035514; PubMed=9371813;
RA Smith D.R., Doucette-Stamm L.A., Deloughery C., Lee H.-M., Dubois J.,
RA Aldredge T., Bashirzadeh R., Blakely D., Cook R., Gilbert K.,
RA Harrison D., Hoang L., Keagle P., Lumm W., Pothier B., Qiu D.,
RA Spadafora R., Vicaire R., Wang Y., Wierzbowski J., Gibson R.,
RA Jiwani N., Caruso A., Bush D., Safer H., Patwell D., Prabhakar S.,
RA McDougall S., Shimer G., Goyal A., Pietrokovski S., Church G.M.,
RA Daniels C.J., Mao J.-I., Rice P., Nolling J., Reeve J.N.;
RT "Complete genome sequence of Methanobacterium thermoautotrophicum
RT deltaH: functional analysis and comparative genomics.";
RL J. Bacteriol. 179:7135-7155(1997).
CC -!- SIMILARITY: BELONGS TO THE ARSR FAMILY OF TRANSCRIPTIONAL
CC REGULATORS.
DR EMBL; AE000908; AAB85626.1; -.
DR InterPro; IPR001845; HTH_ArsR.
DR InterPro; IPR004396; ArsR.
DR Pfam; PF01022; HTH_5; 1.
DR Pfam; PF02800; V4R; 1.
DR SMART; SM00418; HTH_ARSR; 1.
KW Complete proteome; DNA binding; Transcription regulation.
SQ SEQUENCE 251 AA 28229 MW; [illegible] CRC64;

Query Match 74.4%; Score 82; DB 17; Length 251;
Best Local Similarity 66.7%; Pred. No. 64;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FIASSEIG 9
|:| |:|||
Db 7 FIPSSTETG 15

RESULT 8
Q9HPA0
ID Q9HPA0 PRELIMINARY; PRT; 261 AA.
AC Q9HPA0;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE VNG1736C.
GN VNG1736C.
OS Halobacterium sp. (strain NRC-1).
OC Archaea; Euryarchaeota; Halobacteriales; Halobacteriaceae;
OC Halobacterium.
OX NCBI_TaxID=64091;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20564483; PubMed=11016950;
RA Ng W.V., Kennedy S.P., Mahairas G.G., Berquist B., Pan M.,
RA Shukla H.D., Lasky S.R., Baliga N.S., Thorsson V., Sbrogna J.,
RA Swartzell S., Weir D., Hall J., Dahl T.A., Welti R., Goo Y.A.,
RA Leithauser B., Keller K., Cruz R., Danson M.J., Hough D.W.,
RA Maddocks D.G., Jablonski P.E., Krebs M.P., Angevine C.M., Dale H.,
RA Isenbarger T.A., Peck R.F., Pohlschroder M., Spudich J.L., Jung K.-H.,
RA Alam M., Freitas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D.,
RA Ebhardt H., Lowe T.M., Liang P., Riley M., Hood L., DasSarma S.;
RT "Genome sequence of Halobacterium species NRC-1.";
RL Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000).
DR EMBL; AE005079; AAG19970.1; -.
KW Complete proteome.
SQ SEQUENCE 261 AA; 28217 MW; 6F5304AB9E778789 CRC64;

Query Match 74.4%; Score 82; DB 17; Length 261;
Best Local Similarity 66.7%; Pred. No. 65;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 FIASSEIG 9
|:| |:|:|
Db 57 FIIVDSDIG 65

RESULT 9
Q9WTV1
ID Q9WTV1 PRELIMINARY; PRT; 352 AA.
AC Q9WTV1;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE [illegible] HEIR 55 (FRAGMENT).
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN LEWIS;
RA Wendling U., Boots A.M.H., van Eden W.;
RT "Cloning of the rat homologue of Human Cartilage glycoprotein 39 a
RT potential autoantigen in arthritis.";
RL Submitted (APR-1998) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AF062098; AAD22619.1; -.
DR HSSP; P07254; 1CIN.
DR InterPro; IPR000677; 2S_Globulin.
DR InterPro; IPR001223; Glyco_hydro_18.
DR Pfam; PF00704; Glyco_hydro_18; 1.

```
DR   PRINTS; PR00551; 2SGLOBULIN.
FT   NON_TER       1      1
FT   NON_TER     352    352
SQ   SEQUENCE   352 AA;  39391 MW;  CBDE991610AC936C CRC64;


  Query Match             74.4%;  Score 32;  DB 11;  Length 352;
  Best Local Similarity   100.0%;  Pred. No. 90;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 FTLASSE 7
           |||||||
Db     244 FTLASSE 250


RESULT  10
Q90Y50
ID   Q90Y50      PRELIMINARY;      PRT;   372 AA.
AC   Q90Y50;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   COXSACKIEVIRUS AND ADENOVIRUS RECEPTOR LIKE PROTEIN.
GN   CAR.
OS   Brachydanio rerio (Zebrafish) (Zebra danio).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;
OC   Cypriniformes; Cyprinidae; Danio.
OX   NCBI_TaxID=7955;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE 20535037; PubMed 11080637;
RA   van Raaij M.J., Chouin E., van der Zandt H., Bergelson J.M.,
RA   Cusack S.;
RT   "Dimeric structure of the coxsackievirus and adenovirus receptor D1
RT   domain at 1.7 A resolution.";
RL   Structure 8:1147-1155(2000).
DR   EMBL; AF268197; AAK58592.1; -.
KW   Receptor.
SQ   SEQUENCE   372 AA;  40664 MW;  C363B71E7601C74A CRC64;


  Query Match             74.4%;  Score 32;  DB 13;  Length 372;
  Best Local Similarity   66.7%;  Pred. No. 95;
  Matches    6;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 FTLASSETG 9
           |||||  :||
Db      47 FTLASDDSG 55


RESULT  11
Q99J84
ID   Q99J84      PRELIMINARY;      PRT;   381 AA.
AC   Q99J84;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-OCT-2001 (TrEMBLrel. 18, Last annotation update)
DE   SIMILAR TO CHITINASE 3 LIKE 1 (CARTILAGE GLYCOPROTEIN 39).
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Strausberg R.;
RL   Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Strausberg R.;
RL   Submitted (FEB-2001) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RA   Strausberg R.;
RL   Submitted (MAR-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC005611; AAH05611.1; -.
DR   EMBL; BC004780; AAH04780.1; -.
DR   EMBL; BC004734; AAH04734.1; -.
DR   InterPro; IPR001223; Glyco_hydro_18.
DR   Pfam; PF00704; Glyco_hydro_18; 1.
SQ   SEQUENCE   381 AA;  42979 MW;  EE6588C5AE9D4450 CRC64;


  Query Match             74.4%;  Score 32;  DB 11;  Length 381;
  Best Local Similarity   100.0%;  Pred. No. 97;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 FTLASSE 7
           |||||||
Db     266 FTLASSE 272


RESULT  12
Q9SXN2
ID   Q9SXN2      PRELIMINARY;      PRT;   446 AA.
AC   Q9SXN2;
DT   01-MAY-2000 (TrEMBLrel. 14, Created)
DT   01-MAY-2000 (TrEMBLrel. 14, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   1-AMINOCYCLOPROPANE-1-CARBOXYLATE SYNTHASE.
GN   PPPACS2.
OS   Pyrus pyrifolia (Japanese pear) (Pyrus serotina).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids I; Rosales; Rosaceae; Maloideae; Pyrus.
OX   NCBI_TaxID=4767;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=KIKUSUI; TISSUE=FRUIT;
RX   MEDLINE 99168766; PubMed 10071208;
RA   Itai A., Kawata T., Tanabe K., Tamura F., Uchiyama M., Tomomitsu M.,
RA   Shiraiwa N.;
RT   "Identification of 1-aminocyclopropane-1-carboxylic acid synthase
RT   genes controlling the ethylene level of ripening fruit in Japanese
RT   pear (Pyrus pyrifolia Nakai).";
RL   Mol. Gen. Genet. 261:42-49(1999).
CC   -!- COFACTOR: PYRIDOXAL PHOSPHATE (BY SIMILARITY).
CC   -!- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC   -!- MISCELLANEOUS: IN EUKARYOTES THERE ARE TWO ISOZYMES: A CYTOPLASMIC
CC       ONE AND A MITOCHONDRIAL ONE (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO CLASS I OF PYRIDOXAL PHOSPHATE-DEPENDEN
CC       AMINOTRANSFERASES.
DR   EMBL; AB007639; BAA76388.1; -.
DR   HSSP; P37821; 1B8G.
DR   InterPro; IPR001176; ACC_synthase.
DR   InterPro; IPR001511; Aminotran_1.
DR   Pfam; PF00155; aminotran_1_2; 1.
DR   PRINTS; PR00753; ACCSYNTHASE.
DR   PROSITE; PS00105; AA_TRANSFER_CLASS_1; 1.
KW   Pyridoxal phosphate.
SQ   SEQUENCE   446 AA;  49941 MW;  D0A471970E0D1B05 CRC64;


  Query Match             74.4%;  Score 32;  DB 10;  Length 446;
  Best Local Similarity   87.5%;  Pred. No. 1.2e+02;
  Matches    7;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 FTLASSET 8
           ||| ||||
Db     312 FTLISSET 319


RESULT  13
Q9SJQ2
ID   Q9SJQ2      PRELIMINARY;      PRT;   574 AA.
AC   Q9SJQ2;
```

```
DT   01 MAY 2000 (TrEMBLrel. 13, Created)
DT   01 MAY 2000 (TrEMBLrel. 13, Last sequence update)
DT   01 DEC 2001 (TrEMBLrel. 19, Last annotation update)
DE   AT2G46560 PROTEIN.
GN   AT2G46560.
OS   Arabidopsis thaliana (Mouse ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. COLUMBIA;
RX   MEDLINE=20384867; PubMed=11027297;
RA   Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M. I., Town C.D.,
RA   Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblyum T.V.,
RA   Buell C.R., Ketchum K.A., Lee J.J., Ronning C.M., Koo H., Moffat K.S.,
RA   Cronin L.A., Shen M., VanAken S.E., Umayam L., Tillberg L.J., Cotton C.L.,
RA   Adams M.D., Carrera A.J., Creasy T.H., Goodman H.M., Somerville C.R.,
RA   Copenhaver G.P., Preuss D., Nierman W.C., White O., Eisen J.A.,
RA   Salzberg S.L., Fraser C.M., Venter J.C.;
RT   "Sequence and analysis of chromosome 2 of the plant Arabidopsis
RT   thaliana.";
RL   Nature 402:761-768(1999).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. COLUMBIA;
RA   Lin X.;
RL   Submitted (MAR 2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AC006919; AAL24608.1; -.
DR   InterPro; IPR001813; TFIIB.
DR   PROSITE; PS00782; TFIIB; FALSE_NEG.
SQ   SEQUENCE   574 AA;  61191 MW;  07200AF925819604 CRC64;


  Query Match             74.4%;  Score 32;  DB 10;  Length 574;
  Best Local Similarity   87.5%;  Pred. No. 1.5e+02;
  Matches    7;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       2 ILASSEIG 9
           ||:  :||
Db     200 ILASSSIG 207


RESULT  14
Q9N6I8
ID   Q9N6I8      PRELIMINARY;      PRT;  1243 AA.
AC   Q9N6I8;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   DKFZp434L0425.
GN   Leishmania major.
OS   Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
OX   NCBI_TaxID=5664;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=FRIEDLIN;
RX   MEDLINE=99280440; PubMed=10352176;
RA   Ravel C., Dubessay P., Brette C., Blaineau C., Bastien P., Pages M.;
RT   "High conservation of the timescale organisation of chromosome 5
RT   between two pathogenic leishmania species.";
RL   Nucleic Acids Res. 27:2473-2477(1999).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=FRIEDLIN;
RA   Ravel C., Dubessay P., Blaineau C., Lignon M.-F., Bastien P.,
RA   Crobu L., Lebrevre M., Pages M.;
RL   Submitted (FEB 2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF239995; AAF68679.1; -.
DR   InterPro; IPR015171; Amiloride_1.
DR   InterPro; IPR003892; Phosphopant_attach.
DR   InterPro; IPR003954; RRM.
DR   InterPro; IPR003955; RRM_2.
DR   SMART; SM00362; RRM_2; 2.
DR   PROSITE; PS50102; RRM; 1.
SQ   SEQUENCE   1243 AA;  142789 MW;  4A9659B3949EB23E CRC64;


  Query Match             74.4%;  Score 32;  DB 5;  Length 1,243;
  Best Local Similarity   66.7%;  Pred. No. 3.4e+02;
  Matches    6;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 FILASSEIG 9
           :| |:|:|:|
Db    1178 FTAAASQIG 1186


RESULT  15
Q94MX7
ID   Q94MX7      PRELIMINARY;      PRT;  186 AA.
AC   Q94MX7;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   ORF44.
GN   ORF44.
OS   Haemophilus influenzae phage HP2.
OC   Viruses.
OX   NCBI_TaxID=157239;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Williams B.J., Golomb M., Olson M.V., Smith A.L.;
RT   "The HP2 genome of Haemophilus influenzae.";
RL   Submitted (FEB 2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY027935; AAK97818.1; -.
SQ   SEQUENCE   186 AA;  20615 MW;  77A25586ZBFAEF3A CRC64;


  Query Match             72.1%;  Score 31;  DB 9;  Length 186;
  Best Local Similarity   66.7%;  Pred. No. 7e;
  Matches    6;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 FILASSEIG 9
           | ||  :||
Db     72 FNLAEAEFG 80


Search completed: August 26, 2002, 06:28:05
Job time: 1031 sec
```

Geneseq version 4.5
Copyright (c) 2002 - 2002 Compugen Ltd.

SW protein search, using SW model

Run on: August 26, 2002, 06:13:24 ; Search time 139.28 Seconds
(without alignments)
7,127 Million cell updates/sec

Title: US-09-744-282-2
Perfect score: 43
Sequence: 1 EILASSEIG 9

Scoring table: BC SUM62
Gapop 10.0 , Gapext 0.5

Searched: 74,574 seqs, 11,043,796 residues

Total number of hits satisfying chosen parameters: 747,574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_032002.pep:*
1: /cgn2_6/ptodata/1/bl_pep/A_Geneseq_032002.pep:*
2: /cgn2_6/ptodata/1/bl_pep/A_Geneseq_032002.pep:*
3: /cgn2_6/ptodata/1/bl_pep/A_Geneseq_032002.pep:*
4: /SIDS1/gcgdata/bl_pep/A_Geneseq_032002.pep:*
5: /cgn2_6/ptodata/1/bl_pep/A_Geneseq_032002.pep:*
6: /cgn2_6/ptodata/1/bl_pep/A_Geneseq_032002.pep:*
7: /cgn2_6/ptodata/1/bl_pep/A_Geneseq_032002.pep:*
8: /SIDS1/gcgdata/bl_pep/A_Geneseq_032002.pep:*
9: /cgn2_6/ptodata/1/bl_pep/A_Geneseq_032002.pep:*
10: /cgn2_6/ptodata/1/bl_pep/A_Geneseq_032002.pep:*
11: /SIDS1/gcgdata/bl_pep/A_Geneseq_032002.pep:*
12: /cgn2_6/ptodata/1/bl_pep/A_Geneseq_032002.pep:*
13: /cgn2_6/ptodata/1/bl_pep/A_Geneseq_032002.pep:*
14: /cgn2_6/ptodata/1/bl_pep/A_Geneseq_032002.pep:*
15: /cgn2_6/ptodata/1/bl_pep/A_Geneseq_032002.pep:*
16: /cgn2_6/ptodata/1/bl_pep/A_Geneseq_032002.pep:*
17: /cgn2_6/ptodata/1/bl_pep/A_Geneseq_032002.pep:*
18: /SIDS1/gcgdata/bl_pep/A_Geneseq_032002.pep:*
19: /cgn2_6/ptodata/1/bl_pep/A_Geneseq_032002.pep:*
20: /cgn2_6/ptodata/1/bl_pep/A_Geneseq_032002.pep:*
21: /cgn2_6/ptodata/1/bl_pep/A_Geneseq_032002.pep:*
22: /cgn2_6/ptodata/1/bl_pep/A_Geneseq_032002.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	43	100.0	9	17	AAR95072	Human cartilage gl
2	43	100.0	9	21	AAY60227	Peptide fragment o
3	43	100.0	13	17	AAR95076	Peptide conta. hC
4	43	100.0	13	19	AAW61597	Human Cartilage gp
5	43	100.0	13	21	AAY60241	Peptide fragment o
6	43	100.0	13	21	AAY77401	Human cartilage gl
7	43	100.0	13	22	AAU03045	Human cartilage gp
8	43	100.0	13	22	AAU03046	Human cartilage gp
9	43	100.0	13	22	AAU03047	Human cartilage gp
10	43	100.0	13	22	AAU03048	Human cartilage gp
11	43	100.0	13	22	AAU03054	Human cartilage gp
12	43	100.0	13	22	AAU03055	Human cartilage gp
13	43	100.0	13	22	AAU03057	Human cartilage gp
14	43	100.0	13	22	AAU03058	Human cartilage gp
15	43	100.0	13	22	AAU03059	Human cartilage gp
16	43	100.0	13	22	AAU03062	Human cartilage gp
17	43	100.0	13	22	AAU03064	Human cartilage gp
18	43	100.0	13	22	AAU03065	Human cartilage gp
19	43	100.0	13	22	AAU03066	Human cartilage gp
20	43	100.0	13	22	AAU03067	Human cartilage gp
21	43	100.0	13	22	AAU03072	Human cartilage gp
22	43	100.0	13	22	AAU03073	Human cartilage gp
23	43	100.0	14	22	AAU03070	Human cartilage gp
24	43	100.0	14	22	AAU03074	Human cartilage gp
25	43	100.0	15	19	AAW61598	Human Cartilage gp
26	43	100.0	16	18	AAW45140	Human cartilage gl
27	43	100.0	16	19	AAW61599	Human Cartilage gp
28	43	100.0	105	22	AAU02074	Synthetic human ta
29	43	100.0	199	22	ABB56452	Human, single chain
30	43	100.0	199	22	ABB56456	Human, single chain
31	43	100.0	208	22	ABB56475	Human, single chain
32	43	100.0	209	22	ABB56454	Human, single chain
33	43	100.0	210	22	ABB56442	Human, single chain
34	43	100.0	211	22	ABB56441	Human, single chain
35	43	100.0	211	22	ABB56455	Human, single chain
36	43	100.0	212	22	ABB56439	Human, single chain
37	43	100.0	212	22	ABB56440	Human, single chain
38	43	100.0	212	22	ABB56477	Human, single chain
39	43	100.0	213	22	ABB56443	Human, single chain
40	43	100.0	213	22	ABB56445	Human, single chain
41	43	100.0	213	22	ABB56446	Human, single chain
42	43	100.0	213	22	ABB56453	Human, single chain
43	43	100.0	214	22	ABB56444	Human single chain
44	43	100.0	216	22	ABB56447	Human single chain
45	43	100.0	216	22	ABB56448	Human single chain

ALIGNMENTS

RESULT 1
AAR95072
ID AAR95072 standard; peptide; 9 AA.
XX
AC AAR95072;
XX
DT 17-DEC-1996 (first entry)
XX
DE Human cartilage glycoprotein 39 autoantigen, derived peptide (11).
XX
KW Human; cartilage; glycoprotein 39; hC gp-39; autoantigen;
KW Indication; T cell; tolerance induced; abstraction;
KW rheumatoid arthritis, disease model; diagnosis.
XX
OS Homo sapiens.
XX
PN WO9613517-A1.
XX
PD 09-MAY-1996.
XX
PF 25-OCT-1995; 95WO-EP04201.
XX
PR 07-APR-1995; 95EP-0200886.
PR 27-OCT-1994; 94EP-0203128.
XX
PA (AKZO) AKZO NOBEL NV.
XX
PI Boots AMH, Verheijden GFM;
XX
DR WPI; 1996-239447/24.
XX
PT Peptide(s) derived from human cartilage gp-39 autoantigen used
PT to induce specific T cell tolerance, esp. for treatment of
PT rheumatoid arthritis, also useful in producing arthritic animal

```
PT   models
XX
PS   Claim 1; Page 28; 37pp; English.
XX
CC   The present peptide, derived from the human cartilage glycoprotein
CC   39 (hC gp 39) autoantigen, can be used to induce specific T-cell
CC   tolerance to hC gp 39 in a patient with T-cell mediated cartilage
CC   destruction, esp. rheumatoid arthritis (RA). It may also be used to
CC   induce RA in mice, etc. (human disease models), and to diagnose the
CC   presence of activated T-cells autoreactive with hC gp 39, i.e., to
CC   identify patients suitable for tolerising treatment.
CC   Therapeutically the peptide is given in a dose of 0.1-100 micromol/kg
CC   by injection, or 10-50 microg/kg to produce a disease model. The
CC   peptide has a specific tolerising effect on autoreactive T-cells,
CC   but leaves other components of the immune system intact, and
CC   therefore should have few side effects. In an assay to determine
CC   the peptide's binding to HLA DR mols., specifically DR1*01 (known
CC   to be associated with an increased risk of RA development), the
CC   peptide gave a DCT blot IC(50) (micromolar) value of 0.008 compared
CC   to >1.5 for the control peptide HA(307-319)F.
XX
SQ   Sequence   9 AA;


  Query Match             100.0%;  Score 43;  DB 17;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 6.4e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 FTLASSETG 9
             |||||||||
Db       1 FTLASSETG 9


RESULT   2
AAY69227
ID   AAY69227 standard; peptide; 9 AA.
XX
AC   AAY69227;
XX
DT   30-MAY-2000  (first entry)
XX
DE   Peptide fragment of human cartilage gp-39 (HC gp-39) protein.
XX
KW   Human cartilage gp-39 protein; HC gp-39 protein; T-cell activator;
KW   inflammatory disease; lymphocyte; antigen; bystander suppression;
KW   antigen specific T-cell tolerance; autoimmune disorder;
KW   inflammatory disease; rheumatoid arthritis;
KW   anti-inflammatory disease; autoantigen.
XX
OS   Homo sapiens.
XX
PN   WO2000004137-A2.
XX
PD   03-FEB-2000.
XX
PF   19-JUL-1999;   99WO-EP05451.
XX
PR   23-JUL-1998;   98EP-0202471.
XX
PA   (AKZO ) AKZO NOBEL NV.
XX
PI   Millenburg AMM,  Boots AMH;
XX
DR   WPI; 2000-182537/16.
XX
PT   Use of cartilage protein or peptides for treating inflammatory diseases
PT   especially rheumatoid arthritis -
XX
PS   Claim 2; Page 17; 20pp; English.
XX
CC   AAY69226-34 represent fragments of the human cartilage gp-39 (HC gp-39)
CC   protein. The peptides are T-cell activators. The peptides are used for
CC   preventing inflammatory disease, by modulating lymphocyte reactive
CC   to antigens other than HC gp-39, present in the same tissue -
CC   HC gp-39. By the induction of antigen specific T-cell tolerance
CC   autoimmune disorders can be treated by bystander suppression. The
CC   HC gp-39 peptides are useful for preventing inflammatory disease
CC   especially rheumatoid arthritis. The fragments may also be used
CC   treating autoimmune inflammatory diseases. The HC gp-39 peptides are
CC   especially beneficial in autoimmune conditions in which the
CC   antigens associated with the condition are unknown.
XX
SQ   Sequence   9 AA


  Query Match             100.0%;  Score 43;  DB 21;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 6.4e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 FTLASSETG 9
             |||||||||
Db       1 FTLASSETG 9


RESULT   3
AAR95076
ID   AAR95076 standard; peptide; 13 AA.
XX
AC   AAR95076;
XX
DT   17-DEC-1996  (first entry)
XX
DE   Peptide conta. hC gp-39 autoantigen derived peptide (II).
XX
KW   Human; cartilage; glycoprotein 39; hC gp-39; autoantigen;
KW   induction; T cell; tolerance; mediated; destruction;
KW   rheumatoid arthritis; disease model; diagnosis.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Peptide          3..11
FT                    /note= "HC gp-39 peptide (II)"
XX
PN   WO9613517-A1.
XX
PD   09-MAY-1996.
XX
PF   25-OCT-1995;    95WO-EP04105.
XX
PR   27-OCT-1994;    94EP-0203128.
XX
PA   (AKZO ) AKZO NOBEL NV.
XX
PI   Boots AMH,  Verbeijden CFM;
XX
DR   WPI; 1996-244774/24.
XX
CT   Peptide(s) derived from human cartilage gp-39 autoantigen, used
PT   to induce specific T cell tolerance, esp. for treatment of
PT   rheumatoid arthritis, also useful in producing arthritic animal
PT   models
XX
PS   Claim 2; Page 28; 37pp; English.
XX
CC   The present peptide, contg. a peptide derived from the human
CC   cartilage glycoprotein 39 (hC gp-39) autoantigen, can be used to
CC   induce specific T-cell tolerance to hC gp-39 in a patient with
CC   T-cell mediated cartilage destruction, esp. rheumatoid arthritis
CC   (RA). It may also be used to induce RA in mice, etc. (human disease
CC   models), and to diagnose the presence of activated T-cells
CC   autoreactive with hC gp-39, i.e., to identify patients suitable for
CC   tolerising treatment.
CC   Therapeutically the peptide is given in a dose of 0.1-100 micromol/kg
CC   by injection, or 10-50 microg/kg to produce a disease model. The
```

```
CC   peptide has a specific tolerising effect on autoreactive T cells,
CC   but leaves other components of the immune system intact, and
CC   therefore should have few side effects.
XX
SQ   Sequence   13 AA;


  Query Match             1 0.0%;  Score 43;  DB 17;  Length 13;
  Best Local Similarity   1 0.0%;  Pred. No. 0.00059;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 FTLASSETG 9
           ||| ||| |
Db       3 ftlassetq 11


RESULT   4
AAW61597
ID   AAW61597 standard; protein; 15 AA.
XX
AC   AAW61597;
XX
DT   27-OCT-1998  (first entry)
XX
DE   Human Cartilage gp-39 epitope (1).
XX
KW   Human; cartilage gp-39; monoclonal antibody; cell surface antigen;
KW   rheumatoid arthritis; T cell receptor.
XX
OS   Homo sapiens.
XX
PN   EP805620-A1.
XX
PD   05-AUG-1998.
XX
PF   22-DEC-1997;   97EP-0203769.
XX
PR   27-JUN-1997;   97EP-0201972.
PR   06-DEC-1996;   96EP-0203465.
XX
PA   (ALKU ) AKZO NOBEL NV.
PA   (ALKU ) AKZO NV.
XX
PI   Stoenbakkers EGA;
XX
DR   WPI; 1998-472866/40.
XX
PT   Preparation of monoclonal antibodies against cell surface antigens
PT   used as diagnostic agents and to treat rheumatoid arthritis -
PT   comprises immunising, then enriching, enriching and cloning B cells
PT   and then selecting and cloning hybridoma
XX
PS   Examples; Page 5; 24pp; English.
XX
CC   The human cartilage gp-39 peptides AAW61597-W61599 were used in the
CC   preparation of a monoclonal antibody against a cell surface antigen.
CC   The antibody can be used for treatment of rheumatoid arthritis.  It
CC   is also useful against cell surface antigens and antibodies reactive
CC   with the clonotypic structure of T-cell receptors are used as diagnostic
CC   reagents.  The process allows antibodies to be raised against
CC   antigens which are present in very low amounts or which are
CC   conformationally unstable without the need to screen very large numbers
CC   of hybridomas and without requiring high purity Antigen for enrichment.
XX
SQ   Sequence   13 AA;


  Query Match             1 0.0%;  Score 43;  DB 19;  Length 15;
  Best Local Similarity   1 0.0%;  Pred. No. 0.00059;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 FTLASSETG 9
           | || |||
Db       3 ftlassetq 11


RESULT   5
AAY69231
ID   AAY69231 standard; peptide; 13 AA.
XX
AC   AAY69231;
XX
DT   30-MAY-2000  (first entry)
XX
DE   Peptide fragment of human cartilage gp-39 (HC gp-39) protein.
XX
KW   Human cartilage gp-39 protein; HC gp-39 protein; T-cell activator;
KW   inflammatory disease; lymphocyte; antigen; bystander suppression;
KW   antigen-specific T-cell tolerance; autoimmune disorder;
KW   inflammatory disease; rheumatoid arthritis;
KW   autoimmune inflammatory disease; autoantigen.
XX
OS   Homo sapiens.
XX
PN   WO200004917-A2.
XX
PD   03-FEB-2000.
XX
PF   19-JUL-1999;   99WO-EP05331.
XX
PR   23-JUL-1998;   98EP-0202471.
XX
PA   (ALKU ) AKZO NOBEL NV.
XX
PI   Miltenburg AMM,  Boots AMH;
XX
DR   WPI; 2000-182537/16.
XX
PT   Use of cartilage protein or peptides for treating inflammatory diseases
PT   especially rheumatoid arthritis -
XX
PS   Claim 2; Page 17; 29pp; English.
XX
CC   AAY69226-33 represent fragments of the human cartilage gp-39 (HC gp-39)
CC   protein. The peptides are T cell activators. The peptides are used for
CC   preventing inflammatory diseases, by modulating lymphocyte reactive
CC   to antigens other than HC gp-39, present in the same tissue of
CC   HC gp-39. By the induction of antigen-specific T-cell tolerance,
CC   autoimmune disorders can be treated by bystander suppression. The
CC   HC gp-39 peptides are useful for preventing inflammatory diseases
CC   especially rheumatoid arthritis. The fragments may also be used for
CC   treating autoimmune inflammatory diseases. The HC gp-39 peptides are
CC   especially beneficial in autoimmune conditions in which the
CC   autoantigens associated with the condition are unknown.
XX
SQ   Sequence   13 AA;


  Query Match           100.0%;  Score 43;  DB 21;  Length 13;
  Best Local Similarity  100.0%;  Pred. No. 0.00059;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 FTLASSETG 9
           |||||||||
Db       3 ftlassetq 11


RESULT   6
AAY77401
ID   AAY77401 standard; peptide; 13 AA.
XX
AC   AAY77401;
XX
DT   22-MAY-2000  (first entry)
XX
DE   Human cartilage glycoprotein-39 (HC gp-39), residues 263-275.
```

XX
KW Human cartilage glycoprotein 39; HC gp-39; YKL-40; epitope;
KW autoreactive T cell; tolerisation; desensitisation; autoantigen;
KW rheumatoid arthritis; human chondrocyte protein; YKL 39;
KW bystander suppression; autoimmune disorder; Graves' disease;
KW primary glomerulonephritis; inflammatory bowel disease; arthritis.
XX
OS Homo sapiens.
XX
PN WO200057254-A1.
XX
PD 28-SEP-2000.
XX
PF 16-JUL-1999; 99WO-EP005050.
XX
PR 23-JUL-1998; 98IP-0013247.
XX
PA (ALKU) AKZO NOBEL NV.
XX
PI Verheijden GFM, Boots AMH;
XX
DR WPI; 2000-582644/50.
XX
PT New T cell epitope peptides useful for inducing specific T cell
PT tolerance to an autoantigen for treating autoimmune disorders,
PT especially rheumatoid arthritis
XX
PS Example 1; Page 1a; 26pp; English.
XX
CC The invention relates to peptides constituting amino acids 258-270 and
CC amino acids 266-278 of human chondrocyte protein YKL 39
CC (AAY77399 Y77400, respectively). The peptides of the invention may be
CC used in pharmaceutical compositions for the induction of specific T cell
CC tolerance (desensitisation) to an autoantigen in patients
CC suffering from autoimmune disorders, especially rheumatoid arthritis. A
CC frequently found autoantigen in rheumatoid arthritis sufferers is human
CC cartilage glycoprotein 39 (HC gp 39, also known as YKL-40). Several
CC epitopes of HC gp 39 are recognised by peripheral blood T cells from
CC rheumatoid arthritis patients; one such epitope, HC gp-39 residues
CC 263-275, was recognised by T cells from 8 out of 18 patients tested.
CC YKL 39 shares significant sequence identity to hc gp-39. YKL 39 residues
CC 266-278 (AAY77406) constitutes a mimicry epitope of HC gp 39 (263-275),
CC and may be used for desensitisation of T-cells with reactivity to HC
CC gp-39 (263-275) in rheumatoid arthritis patients. Rheumatoid arthritis
CC may be treated by the systemic administration of compositions comprising
CC the peptides of the invention, optionally in combination with existing
CC treatments. The peptides may also be used for modulating lymphocytes
CC that are reactive to autoantigens other than YKL 39 and HC gp-39 but
CC which are present in the same tissue as these autoantigens. By the
CC induction of antigen specific T-cell tolerance, autoimmune disorders can
CC be treated by bystander suppression. The conditions that may be treated
CC via this approach include Graves' disease, primary glomerulonephritis
CC and inflammatory bowel disease. The peptides, along with a detection
CC agent, may additionally be used in a diagnostic composition. The peptides
CC of the invention have a specific effect on autoreactive T cells whilst
CC leaving other components of the immune system intact. Therefore, the
CC peptides should be safe, having no side effects. The present sequence
CC represents residues 263-275 of human cartilage glycoprotein 39 (HC gp-39)
CC which was used in an exemplification of the present invention.
XX
SQ Sequence 13 AA;

Query Match 100.0%; Score 43; DB 21; Length 13;
Best Local Similarity 100.0%; Pred. No. 2.2e-05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FTLASSETG 9
|||||||||
Db 3 ftlassetq 11

RESULT 7
AAU03045
ID AAU03045 standard; peptide; 13 AA.
XX
AC AAU03045;
XX
DT 12-SEP-2001 (first entry)
XX
DE Human cartilage gp 39 autoimmune peptide #1.
XX
KW Human; autoimmune disease; epitope; antigen; rheumatoid arthritis;
KW immunosuppression.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Modified site 1
FT /note "Deaminated"
XX
PN WO200129080-A1.
XX
PD 26-APR-2001.
XX
PF 12-OCT-2000; 2000WO-EP010230.
XX
PR 18-OCT-1999; 99EP-0203427.
XX
PA (ALKU) AKZO NOBEL NV.
XX
PI Van Stavoren CJ, Timmers CM, Van Galen PJM, Knegtel RMA, Boots AMH,
PI Miltenburg AMM;
XX
DR WPI; 2001-367296/38.
XX
PT Modified peptides derived from human cartilage gp 39 protein, useful
PT are useful in immunotherapy, especially in the treatment of arthritis
PT
XX
PS Claim 1; Page 36; 42pp; English.
XX
CC The sequence represents Human cartilage gp 39 protein epitope (amino
CC acids 263-275) which is used as the template for making modified
CC peptides for use in immunotherapy. The peptides are useful for inducing
CC tolerance induction in patients suffering from autoimmune diseases such
CC as rheumatoid arthritis. The peptides provide antigen specific, non-toxic
CC immunosuppression therapy as a very attractive alternative to
CC non-specific immunosuppression.
XX
SQ Sequence 13 AA;

Query Match 100.0%; Score 43; DB 21; Length 13;
Best Local Similarity 100.0%; Pred. No. 2.2e-05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FTLASSETG 9
|||||||||
Db 3 ftlassetq 11

RESULT 8
AAU03046
ID AAU03046 standard; peptide; 13 AA.
XX
AC AAU03046;
XX
DT 12-SEP-2001 (first entry)
XX
DE Human cartilage gp 39 autoimmune peptide #2.
XX
KW human; autoimmune disease; epitope; antigen; rheumatoid arthritis;
KW immunosuppression.
XX
OS Homo sapiens.

OS Synthetic.
XX
FH Key Location/Qualifiers
FT Modified site 1
FT /label= OTHER
FT /note= "other= desaminoarginiinyl"
FT Modified site 13
FT /note= "C-terminus is amidated"
XX
PN WO200129081 A1.
XX
PD 26-APR-2001.
XX
PF 12-OCT-2000; 2000WO-EP10230.
XX
PR 18-OCT 1999; 99EP-0203427.
XX
PA (ALKO) AKZO NOBEL NV.
XX
PI Van Staveren CJ, Timmers CM, Van Galen PJM, Kneetel RMA, Boots AMH,
PI Miltenburg AMM
XX
DR WPI; 2001-367293/38.
XX
PT Modified peptides derived from human cartilage gp-39 protein epitope
PT are useful in immun therapy, especially in the treatment of arthritis
PT -
XX
PS Claim 12; Page 39; 52pp; English.
XX
CC The sequence represents a modified peptide derived from Human cartilage
CC gp-39 protein epitope (amino acids 263-275) for use in immunotherapy.
CC The peptides are useful for inducing tolerance induction in patients
CC suffering from autoimmune diseases such as rheumatoid arthritis.
CC The peptides provide antigen-specific, non-toxic immunosuppression
CC therapy as a very attractive alternative to non-specific
CC immunosuppress.on.
XX
SQ Sequence 13 AA;

Query Match 100.0%; Score 43; DB 22; Length 13;
Best Local Similarity 100.0%; Pred. No. 0.0059;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FTLASSETG 9
||||||||| |
Db 3 ttlasseta 11

RESULT 9
AAU03047
ID AAU03047 standard; peptide; 13 AA.
XX
AC AAU03047;
XX
DT 12-SEP-2001 (first entry)
XX
DE Human cartilage gp-39 autoimmune peptide #3.
XX
KW Human; autoimmune disease; epitope; antigen; rheumatoid arthritis;
KW immunosuppression.
XX
OS Homo sapiens.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Modified site 1
FT /label= OTHER
FT /note= "other -desamino-arginiinyl"
XX
PN WO200129081 A1.
XX
PD 26-APR-2001.
XX
PF 12-OCT-2000; 2000WO-EP10230.
XX
PR 18-OCT-1999; 99EP-0203427.
XX
PA (ALKO) AKZO NOBEL NV.
XX
PI Van Staveren CJ, Timmers CM, Van Galen PJM, Kneetel RMA, Boots AMH,
PI Miltenburg AMM;
XX
DR WPI; 2001-367293/38.
XX
PT Modified peptides derived from human cartilage gp-39 protein epitope
PT are useful in immunotherapy, especially in the treatment of arthritis
PT -
XX
PS Claim 12; Page 39; 52pp; English.
XX
CC The sequence represents a modified peptide derived from Human cartilage
CC gp-39 protein epitope (amino acids 263-275) for use in immunotherapy.
CC The peptides are useful for inducing tolerance induction in patients
CC suffering from autoimmune diseases such as rheumatoid arthritis.
CC The peptides provide antigen-specific, non-toxic immunosuppression
CC therapy as a very attractive alternative to non-specific
CC immunosuppression..
XX
SQ Sequence 13 AA;

Query Match 100.0%; Score 43; DB 22; Length 13;
Best Local Similarity 100.0%; Pred. No. 0.0059;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FTLASSETG 9
||||||||| |
Db 3 ttlasseta 11

RESULT 10
AAU03048
ID AAU03048 standard; peptide; 13 AA.
XX
AC AAU03048;
XX
DT 12-SEP-2001 (first entry)
XX
DE Human cartilage gp-39 autoimmune peptide #4.
XX
KW Human; autoimmune disease; epitope; antigen; rheumatoid arthritis;
KW immunosuppression.
XX
OS Homo sapiens.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Modified site 1
FT /note= "Connected to CH3-(CH2CH2)x-CH2-( )"
FT Modified-site 13
FT /note= "C terminus is amidated"
XX
PN WO200129081-A1.
XX
PD 26-APR-2001.
XX
PF 12-OCT-2000; 2000WO-EP10230.
XX
PR 18-OCT-1999; 99EP-0203427.
XX
PA (ALKO) AKZO NOBEL NV.
XX
PI Van Staveren CJ, Timmers CM, Van Galen PJM, Kneetel RMA, Boots AMH,
PI Miltenburg AMM;

```
XX
DR   WPI; 2001-367293/38.
XX
PT   Modified peptides derived from human cartilage gp-39 protein epitope
PT   are useful in immunotherapy, especially in the treatment of arthritis
PT
XX
PS   Claim 12; Page 39; 52pp; English.
XX
CC   The sequence represents a modified peptide derived from Human cartilage
CC   gp-39 protein epitope (amino acids 263-275) for use in immunotherapy.
CC   The peptides are useful for inducing tolerance induction in patients
CC   suffering from autoimmune diseases such as rheumatoid arthritis.
CC   The peptides provide antigen specific, non toxic immunosuppression
CC   therapy as a very attractive alternative to non specific
CC   immunosuppression.
XX
SQ   Sequence   13 AA;


  Query Match          100.0%;  Score 43;  DB 22;  Length 13;
  Best Local Similarity  100.0%;  Pred. No. 0.0059;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 FTLASSEIG 9
           |||||||||
Db       3 fflasseig 11


RESULT  11
AAU03054
ID   AAU03054 standard; peptide; 13 AA.
XX
AC   AAU03054;
XX
DT   12-SEP-2001  (first entry)
XX
DE   Human cartilage gp-39 autoimmune peptide #10.
XX
KW   Human; autoimmune disease; epitope; antigen; rheumatoid arthritis;
KW   immunosuppression.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT                   /note= "N-terminal acetyl"
FT   Modified-site   2
FT                   /label= OTHER
FT                   /note= "Other= N[(CH2)2-OH]-CH2-C(=O)"
FT   Modified-site   13
FT                   /note= "C-terminus is amidated"
XX
PN   WO200129081-A1.
XX
PD   26-APR-2001.
XX
PF   12-OCT-2000; 2000WO-EP10230.
XX
PR   18-OCT-1999;   99EP-0203427.
XX
PA   (AKZO ) AKZO NOBEL NV.
XX
PI   Van Staveren CJ,  Timmers CM,  Van Galen PJM,  Kneetel RMA,  Boots AMH;
PI   Miltenburg AMM;
XX
DR   WPI; 2001-367293/38.
XX
PT   Modified peptides derived from human cartilage gp-39 protein epitope
PT   are useful in immunotherapy, especially in the treatment of arthritis
PT
XX
```

```
PS   Claim 12; Page 39; 52pp; English.
XX
CC   The sequence represents a modified peptide derived from Human cartilage
CC   gp-39 protein epitope (amino acids 263-275) for use in immunotherapy.
CC   The peptides are useful for inducing tolerance induction in patients
CC   suffering from autoimmune diseases such as rheumatoid arthritis.
CC   The peptides provide antigen specific, non toxic immunosuppression
CC   therapy as a very attractive alternative to non specific
CC   immunosuppression.
XX
SQ   Sequence   13 AA;


  Query Match          100.0%;  Score 43;  DB 22;  Length 13;
  Best Local Similarity  100.0%;  Pred. No. 0.0059;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 FTLASSEIG 9
           |||||||||
Db       3 fflasseig 11


RESULT  12
AAU03055
ID   AAU03055 standard; peptide; 13 AA.
XX
AC   AAU03055;
XX
DT   12-SEP-2001  (first entry)
XX
DE   Human cartilage gp-39 autoimmune peptide #11.
XX
KW   Human; autoimmune disease; epitope; antigen; rheumatoid arthritis;
KW   immunosuppression.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT                   /note= "N-terminal acetyl"
FT   Modified-site   2
FT                   /label= OTHER
FT                   /note= "Other= N[(CH2)2-OH]-CH2-C(=O)"
FT   Modified-site   12
FT                   /label= OTHER
FT                   /note= "Other= NH-CH(CH(CH3)2)-CH2"
FT   Modified-site   13
FT                   /note= "C-terminus is amidated"
XX
PN   WO200129081-A1.
XX
PD   26-APR-2001.
XX
PF   12-OCT-2000; 2000WO-EP10230.
XX
PR   18-OCT-1999;   99EP-0203427.
XX
PA   (AKZO ) AKZO NOBEL NV.
XX
PI   Van Staveren CJ,  Timmers CM,  Van Galen PJM,  Kneetel RMA,  Boots AMH;
PI   Miltenburg AMM;
XX
DR   WPI; 2001-367293/38.
XX
PT   Modified peptides derived from human cartilage gp-39 protein epitope
PT   are useful in immunotherapy, especially in the treatment of arthritis
PT
XX
PS   Claim 12; Page 39; 52pp; English.
XX
CC   The sequence represents a modified peptide derived from Human cartilage
CC   gp-39 protein epitope (amino acids 263-275) for use in immunotherapy
```

CC The peptides are useful for inducing tolerance induction in patients
CC suffering from autoimmune diseases such as rheumatoid arthritis.
CC The peptides provide antigen-specific, non-toxic immunosuppression
CC therapy as a very attractive alternative to non-specific
CC immunosuppression.
XX
SQ Sequence 13 AA;

Query Match 100.0%; Score 43; DB 22; Length 13;
Best Local Similarity 100.0%; Pred. No. 0.0059;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FILASSEIG 9
||||||||| |
Db 3 FILASSETG 11

RESULT 13
AA003057
ID AA003057 standard; peptide; 13 AA.
XX
AC AA003057;
XX
DT 12-SEP-2001 (first entry)
XX
DE Human cartilage gp-39 autoimmune peptide #13.
XX
KW Human; autoimmune disease; epitope; antigen; rheumatoid arthritis;
KW immunosuppression.
XX
OS Homo sapiens.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Modified site 1
FT /label =OTHER
FT /note "Other- H2N-(CH2)6-C(O)"
XX
PN WO200129081-A1.
XX
PD 26-APR-2001.
XX
PF 12-OCT-2000; 2000WO-EP09830.
XX
PR 18-OCT-1999; 99EP-0203427.
XX
PA (ALKU) AKZO NOBEL NV.
XX
PI Van Staveren CJ, Timmers CM, Van Galen PJM, Kneetel RMA, Boots AMH;
PI Miltenburg AMM;
XX
DR WPI; 2001-367293/38.
XX
PT Modified peptides derived from human cartilage gp-39 protein epitope
PT are useful in immunotherapy, especially in the treatment of arthritis
PT .
XX
PS Claim 12; Page 49; 52pp; English.
XX
CC The sequence represents a modified peptide derived from Human cartilage
CC gp-39 protein epitope (amino acids 263-275) for use in immunotherapy.
CC The peptides are useful for inducing tolerance induction in patients
CC suffering from autoimmune diseases such as rheumatoid arthritis.
CC The peptides provide antigen-specific, non-toxic immunosuppression
CC therapy as a very attractive alternative to non specific
CC immunosuppression.
XX
SQ Sequence 13 AA;

Query Match 100.0%; Score 43; DB 22; Length 13;
Best Local Similarity 100.0%; Pred. No. 0.0059;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FILASSEIG 9
||||||||| |
Db 3 FILASSETG 11

RESULT 14
AA003058
ID AA003058 standard; peptide; 13 AA.
XX
AC AA003058;
XX
DT 12-SEP-2001 (first entry)
XX
DE Human cartilage gp-39 autoimmune peptide #14.
XX
KW Human; autoimmune disease; epitope; antigen; rheumatoid arthritis;
KW immunosuppression.
XX
OS Homo sapiens.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Modified site 1
FT /label =OTHER
FT /note "Other- H2N-(CH2)6-C(O)"
XX
PN WO200129081-A1.
XX
PD 26-APR-2001.
XX
PF 12-OCT-2000; 2000WO-EP09830.
XX
PR 18-OCT-1999; 99EP-0203427.
XX
PA (ALKU) AKZO NOBEL NV.
XX
PI Van Staveren CJ, Timmers CM, Van Galen PJM, Kneetel RMA, Boots AMH;
PI Miltenburg AMM;
XX
DR WPI; 2001-367293/38.
XX
PT Modified peptides derived from human cartilage gp-39 protein epitope
PT are useful in immunotherapy, especially in the treatment of arthritis
PT .
XX
PS Claim 12, Page 49, 52pp; English.
XX
CC The sequence represents a modified peptide derived from Human cartilage
CC gp-39 protein epitope (amino acids 263-275) for use in immunotherapy.
CC The peptides are useful for inducing tolerance induction in patients
CC suffering from autoimmune diseases such as rheumatoid arthritis.
CC The peptides provide antigen-specific, non-toxic immunosuppression
CC therapy as a very attractive alternative to non specific
CC immunosuppression.
XX
SQ Sequence 13 AA;

Query Match 100.0%; Score 43; DB 22; Length 13;
Best Local Similarity 100.0%; Pred. No. 0.0059;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FILASSEIG 9
||||||||| |
Db 3 FILASSETG 11

RESULT 15
AA003059
ID AA003059 standard; peptide; 13 AA.
XX

```
AC   AAU00359;
XX
DT   12-SEP-2001  (first entry)
XX
DE   Human cartilage gp-39 autoimmune peptide #15.
XX
KW   Human; autoimmune disease; epitope; antigen; rheumatoid arthritis;
KW   immunosuppression.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT                   /note= "N-terminal (N-methyl-nicotinoyl)-"
XX
PN   WO200129081-A1.
XX
PD   26-APR-2001.
XX
PF   12-OCT-2000; 2000WO-EP10230.
XX
PR   18-OCT-1999;  99EP-0203427.
XX
PA   (AKZO ) AKZO NOBEL NV.
XX
PI   Van Staveren CJ, Timmers CM, Van Galen PJM, Kneefel EMA, Boots AMB;
PI   Miltenburg AMM;
XX
DR   WPI; 2001-367293/38.
XX
PT   Modified peptides derived from human cartilage gp-39 protein epitope
PT   are useful in immunotherapy, especially in the treatment of arthritis
PT
XX
PS   Claim 12; Page 39; 52pp; English.
XX
CC   The sequence represents a modified peptide derived from human cartilage
CC   gp-39 protein epitope (amino acids 263-275) for use in immunotherapy.
CC   The peptides are useful for inducing tolerance induction in patients
CC   suffering from autoimmune diseases such as rheumatoid arthritis.
CC   The peptides provide antigen-specific, non-toxic immunosuppression
CC   therapy as a very attractive alternative to non-specific
CC   immunosuppression.
XX
SQ   Sequence   13 AA;


  Query Match          100.0%;  Score 43;  DB 22;  Length 13;
  Best Local Similarity  100.0%;  Pred. No. 0.0059;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 FTLASSETG 9
           |||||||||
Db       8 fTlAsSeTg 16



Search completed: August 26, 2002, 06:13:24
Job time: 266 sec
```

GenCore version 4.5
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 26, 2002, 05:23:25 ; Search time 40.22 Seconds
(without alignments)
4.756 Million cell updates/sec

Title: US-09-744-282-2
Perfect score: 48
Sequence: 1 FTLASSETG 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0
Maximum DB seq length: 2100000 0

Post-processing: Min match 0%
Maximum Match 1.0%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Prob. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	48	100.0	9	1	US-08-619-645-2	Sequence 2, Appli
2	48	100.0	9	2	US-08-644-493-2	Sequence 2, Appli
3	48	100.0	18	1	US-08-619-645-6	Sequence 6, Appli
4	48	100.0	18	2	US-08-644-493-6	Sequence 6, Appli
5	48	100.0	16	4	US-09-171-705-41	Sequence 41, Appl
6	44	79.1	385	2	US-08-694-915-2	Sequence 2, Appli
7	44	79.1	416	2	US-08-694-915-4	Sequence 4, Appli
8	43	76.7	125	1	US-08-478-039-69	Sequence 69, Appl
9	43	76.7	125	1	US-08-476-349A-69	Sequence 69, Appl
10	42	74.1	18	1	US-08-619-645-5	Sequence 5, Appli
11	42	74.1	18	2	US-08-644-493-5	Sequence 5, Appli
12	42	74.1	16	4	US-09-171-705-40	Sequence 40, Appl
13	42	74.1	502	4	US-09-413-814-106	Sequence 106, App
14	41	72.1	601	2	US-08-606-288-7	Sequence 7, Appli
15	41	72.1	601	2	US-08-606-288-10	Sequence 10, Appl
16	41	72.1	601	3	US-09-347-484-7	Sequence 7, Appli
17	41	72.1	601	3	US-09-347-484-10	Sequence 10, Appl
18	40	69.8	265	2	US-08-860-577-4	Sequence 4, Appli
19	40	69.8	323	4	US-08-846-826A-2	Sequence 2, Appli
20	40	69.8	424	2	US-08-860-577-8	Sequence 8, Appli
21	29	67.4	178	3	US-09-600-630-23	Sequence 23, Appl
22	29	67.4	178	3	US-08-862-730C-23	Sequence 23, Appl
23	29	67.4	178	4	US-09-417-455-10	Sequence 10, Appl
24	29	67.4	178	4	US-09-348-942-10	Sequence 10, Appl
25	29	67.4	178	4	US-09-416-081-6	Sequence 6, Appli
26	29	67.4	401	2	US-08-561-07A-4	Sequence 4, Appli
27	29	67.4	401	2	US-08-561-07A-6	Sequence 6, Appli
28	29	67.4	642	2	US-08-245-511-48	Sequence 48, Appl
29	29	67.4	642	2	US-08-600-993A-48	Sequence 48, Appl
30	29	67.4	2254	2	US-08-286-819A-28	Sequence 28, Appl
31	29	67.4	2254	3	US-08-980-357-28	Sequence 28, Appl
32	28	65.1	16	4	US-09-171-705-42	Sequence 42, Appl
33	28	65.1	332	2	US-08-671-978A-8	Sequence 8, Appli
34	28	65.1	527	2	US-08-522-229B-2	Sequence 2, Appli
35	28	65.1	527	2	US-09-102-433-2	Sequence 2, Appli
36	28	65.1	1588	1	US-08-698-551-16	Sequence 16, Appl
37	28	65.1	1588	2	US-08-602-226-16	Sequence 16, Appl
38	28	65.1	1588	2	US-08-839-032A-16	Sequence 16, Appl
39	27	62.8	261	4	US-08-768-373-2	Sequence 2, Appli
40	27	62.8	295	2	US-08-542-003-4	Sequence 4, Appli
41	27	62.8	295	2	US-08-322-760A-4	Sequence 4, Appli
42	27	62.8	361	4	US-08-838-151A-2	Sequence 2, Appli
43	27	62.8	361	4	US-08-838-151A-4	Sequence 4, Appli
44	27	62.8	361	4	US-08-838-151A-6	Sequence 6, Appli
45	27	62.8	361	4	US-08-838-151A-8	Sequence 8, Appli

ALIGNMENTS

RESULT 1
US-08-619-645-2
; Sequence 2, Application US/08619645
; Patent No. 5736507
; GENERAL INFORMATION:
; APPLICANT: Boots, Anna M.H.
; APPLICANT: Verheijden, Gijsbertus F.M.
; TITLE OF INVENTION: New, cyclic peptides derived from
; TITLE OF INVENTION: and analogs for use in immunotherapy of autoimmune
; TITLE OF INVENTION: disease
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Akzo No. 5736507el Patent Department
; STREET: 1300 Piccard Drive, Suite 206
; CITY: Rockville
; STATE: Maryland
; COUNTRY: U.S.A.
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25(EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/619,645
; FILING DATE: 25-MAR-1996
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Gormley, Mary E.
; REGISTRATION NUMBER: 34,409
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 947-4433
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-619-645-2

Query Match 100.0%; Score 48; DB 1; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.7e-05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FTLASSETG 9
||||||||| |
Db 1 FTLASSETG 9

```
RESULT   2
US-08-634-493-2
; Sequence 2, Application US/08634493
; Patent No. 5843449
;  GENERAL INFORMATION:
;     APPLICANT:  A.M.H. Boots
;     APPLICANT:  G.F.M. Verheijden
;     APPLICANT:  E.S. Bos
;     TITLE OF INVENTION:  No. 5843449el Peptides derivod from autoantigen for use
;     TITLE OF INVENTION:  in Immunotherapy of Autoimmune Diseases
;     NUMBER OF SEQUENCES:  10
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Akzo No. 5843449el Patent Department
;       STREET:  1300 Piccard Drive, Suite 206
;       CITY:  Rockville
;       STATE:  Maryland
;       COUNTRY:  USA
;       ZIP:  20850
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/634,493
;       FILING DATE:  18 APR-1996
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/619,645
;       FILING DATE:  25 MAR-1996
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  PCT/EP95/04201
;       FILING DATE:  25-OCT 1995
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  NL 9420l287
;       FILING DATE:  27 OCT 1994
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  NL 95200886O
;       FILING DATE:  07-APR 1995
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (30l) 258-5200
;       TELEFAX:  (30l) 977-0847
;   INFORMATION FOR SEQ ID NO:  2:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  9 amino acids
;       TYPE:  amino acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  peptide
US-08-634-493-2


  Query Match             100.0%;  Score 43;  DB 2;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 1.7e-05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 FILASSEIG 9
           |||||||||
Db       1 FILASSEIG 9


RESULT   3
US-08-619-645-6
; Sequence 6, Application US/08619645
; Patent No. 5736507
;  GENERAL INFORMATION:
;     APPLICANT:  Boots, Anna M.H.
;     APPLICANT:  Verheijden, Gijsbertus F.M.
;     TITLE OF INVENTION:  No. 5736507el peptides derived from
;     TITLE OF INVENTION:  autoantigen for use in immunotherapy of autoimmune
;     TITLE OF INVENTION:  disease
;     NUMBER OF SEQUENCES:  9
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Akzo No. 5736507el Patent Department
;       STREET:  1300 Piccard Drive, Suite 206
;       CITY:  Rockville
;       STATE:  Maryland
;       COUNTRY:  U.S.A.
;       ZIP:  20850
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC DOS/MS DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25(EPO)
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/619,645
;       FILING DATE:  25-MAR 1996
;       CLASSIFICATION:  514
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Goodlow, Mary E.
;       REGISTRATION NUMBER:  34,409
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (30l) 947-4455
;   INFORMATION FOR SEQ ID NO:  6:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  13 amino acids
;       TYPE:  amino acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  peptide
US-08-619-645-6


  Query Match             100.0%;  Score 43;  DB 1;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 1.7e-05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 FILASSEIG 9
           |||||||||
Db       3 FILASSEIG 11


RESULT   4
US-08-634-493-6
; Sequence 6, Application US/08634493
; Patent No. 5843449
;  GENERAL INFORMATION:
;     APPLICANT:  A.M.H. Boots
;     APPLICANT:  G.F.M. Verheijden
;     APPLICANT:  E.S. Bos
;     TITLE OF INVENTION:  No. 5843449el Peptides derived from autoantigen for use
;     TITLE OF INVENTION:  in Immunotherapy of Autoimmune Diseases
;     NUMBER OF SEQUENCES:  10
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Akzo No. 5843449el Patent Department
;       STREET:  1300 Piccard Drive, Suite 206
;       CITY:  Rockville
;       STATE:  Maryland
;       COUNTRY:  USA
;       ZIP:  20850
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/634,493
;       FILING DATE:  18 APR 1996
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/619,645
;       FILING DATE:  25 MAR 1996
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  PCT/EP95/04201
;       FILING DATE:  25-OCT 1995
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  NL 9420l287
```

```
;       FILING DATE:  27-OCT-1994
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  NL 9400860
;       FILING DATE:  07-APR-1994
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (301) 258-5203
;       TELEFAX:  (+1) 427 0847
;   INFORMATION FOR SEQ ID NO:  6:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  12 amino acids
;       TYPE:  amino acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  peptide
US-08-634-493-6


  Query Match             100.0%;  Score 43;  DB 2;  Length 13;
  Best Local Similarity   100.0%;  Pred. No. 0.002;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 FTLASSEIC 9
           |||||||||
Db       4 FTLASSETC 11


RESULT   5
US-09-171-705-41
; Sequence 41, Application US/09171705
; Patent No. 6184204
; GENERAL INFORMATION:
;  APPLICANT: BOOTS, ANNA M.H.
;  APPLICANT: VERHEIJDEN, GILBERTUS F.M.
;  TITLE OF INVENTION: NOVEL PEPTIDES SUITABLE FOR USE IN ANTIGEN SPECIFIC
;  TITLE OF INVENTION: IMMUNOSUPPRESSIVE THERAPY
;  FILE REFERENCE: 7/96198 US
;  CURRENT APPLICATION NUMBER: US/09/171,705
;  CURRENT FILING DATE: 1999-02-09
;  NUMBER OF SEQ ID NOS: 78
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 41
;   LENGTH: 16
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: DERIVED FROM
;   OTHER INFORMATION: SEQUENCE OF HUMAN CARTILAGE (HC)  39 KDALTIN
US-09-171-705-41


  Query Match             100.0%;  Score 43;  DB 4;  Length 16;
  Best Local Similarity   100.0%;  Pred. No. 0.0025;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 FTLASSEIC 9
           |||||||||
Db       4 FTLASSETC 12


RESULT   6
US-08-694-915-2
; Sequence 2, Application US/08694915
; Patent No. 5811535
;  GENERAL INFORMATION:
;    APPLICANT:  Adamou, Julie
;    APPLICANT:  Kirkpatrick, Robert
;    APPLICANT:  Rosenberg, Martin
;    TITLE OF INVENTION:  HUMAN CARTILAGE GP39 LIKE GENE
;    NUMBER OF SEQUENCES:  7
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SmithKline Beecham Corporation
;      STREET:  709 Swedeland Road
;      CITY:  King of Prussia
;      STATE:  PA
;      COUNTRY:  USA
;      ZIP:  19406-2799
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ Version 1.5
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/694,915
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Han, William T
;      REGISTRATION NUMBER:  34,344
;      REFERENCE/DOCKET NUMBER:  ATG50017
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  610-270-5219
;      TELEFAX:  610-270-5090
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  385 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    FRAGMENT TYPE:  N-terminal
;    ORIGINAL SOURCE:
US-08-694-915-2


  Query Match             79.1%;  Score 34;  DB 2;  Length 385;
  Best Local Similarity   87.5%;  Pred. No. 11;
  Matches    7;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 FTLASSET 8
           ||||:||
Db     268 FTLASAET 275


RESULT   7
US-08-694-915-4
; Sequence 4, Application US/08694915
; Patent No. 5811535
;  GENERAL INFORMATION:
;    APPLICANT:  Adamou, Julie
;    APPLICANT:  Kirkpatrick, Robert
;    APPLICANT:  Rosenberg, Martin
;    TITLE OF INVENTION:  HUMAN CARTILEGE GP39 LIKE GENE
;    NUMBER OF SEQUENCES:  7
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SmithKline Beecham Corporation
;      STREET:  709 Swedeland Road
;      CITY:  King of Prussia
;      STATE:  PA
;      COUNTRY:  USA
;      ZIP:  19406-2799
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ Version 1.5
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/694,915
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
```

```
;       APPLICATION NUMBER:
;       FILING DATE:
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Ban, William T
;       REGISTRATION NUMBER:  34,344
;       REFERENCE/DOCKET NUMBER:  ATG50017
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  610-270-5219
;       TELEFAX:  610-270-5090
;       TELEX:
;   INFORMATION FOR SEQ ID NO:  4:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  415 amino acids
;       TYPE:  amino acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  peptide
;     HYPOTHETICAL:  No
;     ANTI-SENSE:  No
;     FRAGMENT TYPE:  N-terminal
;     ORIGINAL SOURCE:
US-08-604-015-4


  Query Match             79.1%;  Score 34;  DB 2;  Length 415;
  Best Local Similarity   87.5%;  Pred. No. 12;
  Matches    7;  Conservative    1;  Mismatches    0;  Indels    1;  Gaps    0;

Qy       1 FILASSEI 8
           |:|||:||
Db     209 FILASAET 306


RESULT   8
US-08-478-039-69
; Sequence 69, Application US/08478039
; Patent No. 5681722
;  GENERAL INFORMATION:
;     APPLICANT:  Newman, Roland A.
;     APPLICANT:  Hanna, Nabil
;     APPLICANT:  Raab, Ronald W.
;     TITLE OF INVENTION:  Recombinant Antibodies for Human Therapy
;     NUMBER OF SEQUENCES:  114
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  BURNS, DOANE, SWECKER & MATHIS
;       STREET:  699 Prince St.
;       CITY:  Alexandria
;       STATE:  VA
;       COUNTRY:  USA
;       ZIP:  22313-1404
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.30
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/478,039
;       FILING DATE:  07-JUN-1995
;       CLASSIFICATION:  435
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/479,072
;       FILING DATE:  25-JAN-1995
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 07/912,292
;       FILING DATE:  10-JUL-1992
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 07/856,281
;       FILING DATE:  23-MAR-1992
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 07/735,064
;       FILING DATE:  25-JUL-1991
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Toskin Esq., Robin L.
;       REGISTRATION NUMBER:  35,000
;       REFERENCE/DOCKET NUMBER:  01272.100
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  703-836-6620
;       TELEFAX:  703-836-2021
;   INFORMATION FOR SEQ ID NO:  69:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  425 amino acids
;       TYPE:  amino acid
;       STRANDEDNESS:  not relevant
;       TOPOLOGY:  not relevant
;     MOLECULE TYPE:  peptide
;     ORIGINAL SOURCE:
;       ORGANISM:  Monkey
;     POSITION IN GENOME:
;       CHROMOSOME/SEGMENT:  VH2 clone 2-10
US-08-478-039-69


  Query Match             79.1%;  Score 34;  DB 1;  Length 425;
  Best Local Similarity   66.7%;  Pred. No. 4.9;
  Matches    6;  Conservative    3;  Mismatches    0;  Indels    1;  Gaps    1;

Qy       1 FILASSEIG 9
           :|:|:|:|:
Db      27 FSLSISEIG 35


RESULT   9
US-08-476-349A-69
; Sequence 69, Application US/08476349A
; Patent No. 5756105
;  GENERAL INFORMATION:
;     APPLICANT:  Newman, Roland A.
;     APPLICANT:  Hanna, Nabil
;     APPLICANT:  Raab, Ronald W.
;     TITLE OF INVENTION:  Recombinant Antibodies for Human Therapy
;     NUMBER OF SEQUENCES:  114
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  BURNS, DOANE, SWECKER & MATHIS
;       STREET:  699 Prince St.
;       CITY:  Alexandria
;       STATE:  VA
;       COUNTRY:  USA
;       ZIP:  22313-1404
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.30
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/476,349A
;       FILING DATE:  07-JUN-1995
;       CLASSIFICATION:  514
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/479,072
;       FILING DATE:  25-JAN-1995
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 07/912,292
;       FILING DATE:  10-JUL-1992
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 07/856,281
;       FILING DATE:  23-MAR-1992
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 07/735,064
;       FILING DATE:  25-JUL-1991
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Toskin Esq., Robin L.
;       REGISTRATION NUMBER:  35,000
;       REFERENCE/DOCKET NUMBER:  01272.100
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  703-836-6620
;       TELEFAX:  703-836-2021
```

; INFORMATION FOR SEQ ID NO: 69:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 125 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: peptide
; ORIGINAL SOURCE:
; ORGANISM: Monkey
; POSITION IN GENOME:
; CHROMOSOME/SEGMENT: VH2 [illegible] 2-12
US-08-476-349A-69

Query Match 76.7%; Score 33; DB 1; Length 125;
Best Local Similarity 66.7%; Pred. No. 4.8;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FILASSEIG 9
|:|:|||
Db 27 FSLSISEIG 35

RESULT 10
US-08-619-645-5
; Sequence 5, Application US/08619645
; Patent No. 5736507
; GENERAL INFORMATION:
; APPLICANT: Boots, Anna M.H.
; APPLICANT: Verheijden, Gijsbertus F.M.
; TITLE OF INVENTION: Novel peptides derived from
; TITLE OF INVENTION: Antigen for Use in Immunotherapy of Autoimmune
; TITLE OF INVENTION: Diseases
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Akzo No. 5844449el Patent Department
; STREET: 1400 Piccard Drive, Suite 206
; CITY: Rockville
; STATE: Maryland
; COUNTRY: U.S.A.
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25(EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/619,645
; FILING DATE: 25-MAR-1996
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Formey, Mary E.
; REGISTRATION NUMBER: 34,429
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 847-4843
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 13 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-619-645-5

Query Match 74.4%; Score 32; DB 1; Length 13;
Best Local Similarity 100.0%; Pred. No. 0.53;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FILASSE 7
|||||||
Db 7 FILASSE 13

RESULT 11
US-08-634-493-5
; Sequence 5, Application US/08634493
; Patent No. 5844449
; GENERAL INFORMATION:
; APPLICANT: A.M.H. Boots
; APPLICANT: G.F.M. Verheijden
; APPLICANT: E.S. Bos
; TITLE OF INVENTION: No. 5844449el Peptides derived from autoantigen for use
; TITLE OF INVENTION: In Immunotherapy of Autoimmune Diseases
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Akzo No. 5844449el Patent Department
; STREET: 1400 Piccard Drive, Suite 206
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/634,493
; FILING DATE: 18-APR-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/619,645
; FILING DATE: 25-MAR-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/EP95/04201
; FILING DATE: 25-OCT-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: NL 9402031287
; FILING DATE: 27-OCT-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: NL 95200860
; FILING DATE: 07-APR-1995
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 258-5200
; TELEFAX: (301) 977-0847
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 13 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-634-493-5

Query Match 74.4%; Score 32; DB 2; Length 13;
Best Local Similarity 100.0%; Pred. No. 0.53;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FILASSE 7
|||||||
Db 7 FILASSE 13

RESULT 12
US-09-171-705-40
; Sequence 40, Application US/09171705
; Patent No. 6184204
; GENERAL INFORMATION:
; APPLICANT: BOOTS, ANNA M.H.
; APPLICANT: VERHEIJDEN, GILBERTUS F.M.
; TITLE OF INVENTION: NOVEL PEPTIDES SUITABLE FOR USE IN ANTIGEN SPECIFIC
; TITLE OF INVENTION: IMMUNOSUPPRESSIVE THERAPY
; FILE REFERENCE: a/96198 US
; CURRENT APPLICATION NUMBER: US/09/171,705
; CURRENT FILING DATE: 1999-02-09

```
;  NUMBER OF SEQ ID NOS: 78
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 40
;    LENGTH: 16
;    TYPE: PRT
;    ORGANISM: Artificial Sequence
;    FEATURE:
;    OTHER INFORMATION: Description of Artificial Sequence: DERIVED FROM
;    OTHER INFORMATION: SEQUENCE OF HUMAN CARTILAGE GP39   Y-18-CHIN
US-09-171-705-40


  Query Match             74.4%;  Score 32;  DB 4;  Length 16;
  Best Local Similarity   100.0%;  Pred. No. 8.68;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 FILASSE 7
           ||| |||
Db      10 FILASSE 16


RESULT  13
US-09-413-814-106
; Sequence 106, Application US/09413814
; Patent No. 6225064
; GENERAL INFORMATION:
;  APPLICANT: Gesellschaft fuer Biotechnologische Forschung mbH
;  APPLICANT: Bristol-Myers Squibb, Co.
;  APPLICANT: Boyer, Stefan
;  APPLICANT: Bloecker, Helmut
;  APPLICANT: Brandt, Petra
;  APPLICANT: Cino, Earl M
;  APPLICANT: Dougherty, Brian A
;  APPLICANT: Goldberg, Steven L
;  APPLICANT: Hotle, Gerhard
;  APPLICANT: Mueller, Joachim
;  APPLICANT: Reichenbach, Hans
;  TITLE OF INVENTION: DNA Sequences for enzymatic synthesis of polyketide co
;  TITLE OF INVENTION: heteropolyketide compounds
;  FILE REFERENCE: PCT/US 99/23535
;  CURRENT APPLICATION NUMBER: US/09/413,814
;  CURRENT FILING DATE: 1999-10-07
;  EARLIER APPLICATION NUMBER: DE 198 46 466.2
;  EARLIER FILING DATE: 1998-10-09
;  NUMBER OF SEQ ID NOS: 107
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 106
;    LENGTH: 502
;    TYPE: PRT
;    ORGANISM: Sorangium cellulosum
US-09-413-814-106


  Query Match             74.4%;  Score 32;  DB 4;  Length 502;
  Best Local Similarity   77.8%;  Pred. No. 42;
  Matches    7;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 FILASSEIG 9
           | | ||||
Db     182 FVLESSEIG 190


RESULT  14
US-08-606-288-7
; Sequence 7, Application US/08606288
; Patent No. 5955087
;  GENERAL INFORMATION:
;    APPLICANT:  Whittle, N.R.
;    APPLICANT:  Carmichael, J.P.
;    APPLICANT:  Connor, S.E.
;    APPLICANT:  Thompson, H.S.G.
;    APPLICANT:  Wilson, M.J.
;    TITLE OF INVENTION:  Polypeptides Useful as Immunotherapeutic
;    TITLE OF INVENTION:  Agents, and Methods of Polypeptide Preparat
;    NUMBER OF SEQUENCES:  10
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Flehr, Hohbach, Test, Albritton & Herbert
;      STREET:  Suite 3400, Four Embarcadero Center
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  USA
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.2 (EPO)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/606,288
;      FILING DATE:  23-FEB-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  GB 9[illegible]6.7
;      FILING DATE:  24-FEB-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/006,054
;      FILING DATE:  03-JUN-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  GB 9[illegible]478.7
;      FILING DATE:  26-JUL-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Walter H. Dreger
;      REGISTRATION NUMBER:  24,196
;      REFERENCE/DOCKET NUMBER:  A-63294/WHD
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 781-1989
;      TELEFAX:  (415) 398-3249
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  601 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-606-288-7


  Query Match             72.1%;  Score 31;  DB 2;  Length 601;
  Best Local Similarity   75.0%;  Pred. No. 87;
  Matches    6;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 FILASSET 8
           |: ||||
Db     159 FTIISSET 166


RESULT  15
US-08-606-288-10
; Sequence 10, Application US/08606288
; Patent No. 5955087
;  GENERAL INFORMATION:
;    APPLICANT:  Whittle, N.R.
;    APPLICANT:  Carmichael, J.P.
;    APPLICANT:  Connor, S.E.
;    APPLICANT:  Thompson, H.S.G.
;    APPLICANT:  Wilson, M.J.
;    TITLE OF INVENTION:  Polypeptides Useful as Immunotherapeutic
;    TITLE OF INVENTION:  Agents, and Methods of Polypeptide Preparat
;    NUMBER OF SEQUENCES:  10
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Flehr, Hohbach, Test, Albritton & Herbert
;      STREET:  Suite 3400, Four Embarcadero Center
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  USA
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
```

```
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25 (EPO)
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/705,288
;       FILING DATE:  23-FEB-1996
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  GB 9503786.7
;       FILING DATE:  24-FEB-1995
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 60/006034
;       FILING DATE:  08-JUN-1995
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  GB 9515478.7
;       FILING DATE:  28-JUL-1995
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Walter H. Breger
;       REGISTRATION NUMBER:  24,160
;       REFERENCE/DOCKET NUMBER:  A-1284/WHB
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (415) 781-1989
;       TELEFAX:  (415) 398-3249
;   INFORMATION FOR SEQ ID NO:  10:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  601 amino acids
;       TYPE:  amino acid
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  protein
;     HYPOTHETICAL:  N
;     ANTI-SENSE:  NO
US-08-705-288-10

  Query Match             72.1%;  Score 431;  DB 2;  Length 601;
  Best Local Similarity   75.0%;  Pred. No. 8?;
  Matches   ??;  Conservative    ?;  Mismatches    1;  Indels    0;  Gaps    0;

Qy        1 PTLASSEI 8
            ||:   |
Db      159 PTITSSET 166


Search completed: August 26, 2002, 06:23:26
Job time: 897 sec
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 26, 2002, 06:13:45 ; Search time 80.5 Seconds
(without alignments)
10.743 Million cell updates/sec

Title: US-09-744-282-3
Perfect score: 46
Sequence: 1 YDDQESVKS 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_72:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	46	100.0	383	2	A49562	cartilage glycopro
2	43	93.5	384	2	S51427	heparin-binding gl
3	37	80.4	4063	2	T42993	probable spectrin
4	37	80.4	4102	2	T23630	hypothetical prote
5	36	78.3	797	2	S28193	probable DNA-direc
6	36	78.3	5170	2	T15448	hypothetical prote
7	35	76.1	134	2	A45691	hypothetical prote
8	35	76.1	235	2	C00115	cytochrome c type
9	35	76.1	235	2	G07984	cytochrome c-type
10	35	76.1	405	2	S61551	breast regressin
11	34	73.9	441	2	S60618	4 alpha-glucanotra
12	34	73.9	1612	2	S59669	DNA topoisomerase
13	34	73.9	1626	2	A39242	DNA topoisomerase
14	33	71.7	374	2	C01674	hemolysin, probabl
15	33	71.7	534	2	P04835	paraquat-inducible
16	33	71.7	546	2	C90758	paraquat inducible
17	33	71.7	546	2	A85622	paraquat-inducible
18	33	71.7	546	2	A59626	probable secreted
19	33	71.7	827	2	S50714	DNA binding protei
20	33	71.7	1058	2	D08835	CDP1-interacting p
21	33	71.7	1199	2	JC4816	major surface glyc
22	33	71.7	1248	2	I05059	hypothetical prote
23	32	69.6	123	2	E45691	hypothetical prote
24	32	69.6	144	2	A10098	hypothetical prote
25	32	69.6	176	2	E72448	conserved hypothet
26	32	69.6	230	2	I31722	hypothetical prote
27	32	69.6	369	2	JC4292	probable transposa
28	32	69.6	374	2	S15161	long-chain-fatty-a
29	32	69.6	413	2	I43810	methylaspartate am
30	32	69.6	413	2	A99724	3-methylaspartate
31	32	69.6	413	2	A85575	probable methylasp
32	32	69.6	439	2	H81414	probable heat shoc
33	32	69.6	494	1	ALBYAP	alpha-amylase (EC
34	32	69.6	516	2	T00974	probable SF16 prot
35	32	69.6	518	2	T15404	hypothetical prote
36	32	69.6	609	2	I28896	hypothetical protein
37	32	69.6	655	2	A48468	dnaK-type molecula
38	32	69.6	689	2	T40364	hypothetical prote
39	32	69.6	771	2	G71409	probable replicati
40	32	69.6	773	2	S46011	probable pre-mRNA
41	32	69.6	875	2	S75377	probable helicase
42	32	69.6	2946	2	T00867	hypothetical prote
43	31.5	68.5	569	2	T22372	hypothetical prote
44	31	67.4	118	2	B95058	transcription regu
45	31	67.4	118	2	C97927	hypothetical prote

ALIGNMENTS

RESULT 1
A49562
cartilage glycoprotein gp39 precursor - human
N;Alternate names: 39K synovial protein
C;Species: Homo sapiens (man)
C;Date: 23-Mar-1995 #sequence_revision 23-Mar-1995 #text_change 22-Jun-1999
C;Accession: A49562; S10677; A33162
R;Hakala, B.E.; White, C.; Recklies, A.D.
J. Biol. Chem. 268, 25803-25810, 1993
A;Title: Human cartilage gp-39, a major secretory product of articular chondrocytes a
A;Reference number: A49562; MUID:94064658
A;Accession: A49562
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-383 <HAK>
A;Cross-references: GB:M80927; NID:g348911; PIDN:AAA16074.1; PID:g348912
R;Nyirkos, P.; Golds, E.E.
Biochem. J. 269, 265-268, 1990
A;Title: Human synovial cells secrete a 39 kDa protein similar to a bovine mammary pr
A;Reference number: S10677; MUID:90328983
A;Accession: S10677
A;Molecule type: protein
A;Residues: 22-40,'X',42-45 <NY2>
C;Superfamily: Streptomyces chitinase chi40
C;Keywords: cartilage; extracellular protein; glycoprotein
F;1-22/Domain: signal sequence #status predicted <SIG>
F;22-383/Product: cartilage glycoprotein gp39 #status predicted <MAT

Query Match 100.0%; Score 46; DB 2; Length 383;
Best Local Similarity 100.0%; Pred. No. 0.16;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YDDQESVKS 9
||||||||| |
Db 328 YDDQESVKS 336

RESULT 2
S51427
heparin-binding glycoprotein 38K - pig
C;Species: Sus scrofa domestica (domestic pig)
C;Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 22-Jun-1999
C;Accession: S51427
R;Shackelton, L.M.; Mann, D.M.; Millis, A.J.T.
submitted to the EMBL Data Library, January 1995
A;Description: Identification of a 38kDa heparin-binding glycoprotein (gp38k) in diff
A;Reference number: S51427
A;Accession: S51427
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-383 <SHA>

A;Cross-references: EMBL:Z47803; NID:g634097; PIDN:CAA87764.1; PID:g634098
C;Superfamily: Streptomyces chitinase chi40

Query Match 93.5%; Score 43; DB 2; Length 38;
Best Local Similarity 88.9%; Pred. No. 0.64;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YDDQESVKS 9
|||||||||
Db 328 YDDQESVKN 336

RESULT 3
T42998
probable spectrin beta chain - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 11-Jan-2000
C;Accession: T42998
R;Austin, J.; Praitis, V.; McKeown, C.
submitted to the EMBL Data Library, March 1998
A;Description: Sma-1 encodes a BH-spectrin homolog required for C. elegans morphogenesis
A;Reference number: Z22279
A;Accession: T42998
A;Status: preliminary; translated from GB/EMBL/Z92907
A;Molecule type: mRNA
A;Residues: 1-4363 <AUS>
A;Cross-references: EMBL:AF053496; PIDN:AAC08577.1
C;Genetics:
A;Gene: sma-1

Query Match 80.4%; Score 37; DB 2; Length 4363;
Best Local Similarity 66.7%; Pred. No. 1.4e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 YDDQESVKS 9
||| :|:||
Db 4036 YDDSDSIKS 4044

RESULT 4
T24630
hypothetical protein R31.1 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 29-Oct-1999
C;Accession: T24630; T24242
R;Kershaw, J.
submitted to the EMBL Data Library, November 1996
A;Reference number: Z19774
A;Accession: T24630
A;Status: preliminary; translated from GB/EMBL/D08J
A;Molecule type: DNA
A;Residues: 1-4101 <WIL>
A;Cross-references: EMBL:Z81570; PIDN:CAB04608.1; GSPDB:GN00024; CESP:R31.1
A;Experimental source: clone K12G11
R;Lennard, N.
submitted to the EMBL Data Library, July 1996
A;Reference number: Z19862
A;Accession: T24242
A;Status: preliminary; translated from GB/EMBL/Z92E1
A;Molecule type: DNA
A;Residues: 1-4101 <WIL>
A;Cross-references: EMBL:Z75956; PIDN:CAB00140.1; GSPDB:GN00024; CESP:R31.1
A;Experimental source: clone R31
C;Genetics:
A;Gene: CESP:R31.1
A;Map position: 5
A;Introns: 29/3; 99/3; 581/3; 812/3; 961/2; 1151/3; 1254/3; 1...; 1...; 1...; 2...; 2...

Query Match 80.4%; Score 37; DB 2; Length 4101;
Best Local Similarity 66.7%; Prod. No. 1.4e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 YDDQESVKS 9
||| :|:||
Db 4056 YDDSDSIKS 4066

RESULT 5
S28103
probable DNA-directed DNA polymerase (EC 2.7.7.7) - African mosquito-fish (Aphanius [illegible]
C;Species: Aphanius biforquis
C;Date: 28-May-1993 #sequence_revision 28-May-1993 #text_change 22-Jul-1997
C;Accession: S28103
R;Robison, M.M.; Royer, J.C.; Horgen, P.A.
Curr. Genet. 19, 495-502, 1991
A;Title: Homology between mitochondrial DNA of Aphanius [illegible] and B. [illegible]
A;Reference number: S28103; MEDLINE:91347410
A;Accession: S28103
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-797 <Rob>
A;Cross-references: EMBL:Z11515
C;Genetics:
A;Genome: plasmid
C;Keywords: DNA binding, nucleotidase; nucleotidyltransferase

Query Match 78.3%; Score 36; DB 2; Length 797;
Best Local Similarity 75.0%; Pred. No. 6;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 YDDQESVK 8
|||| |:|
Db 400 YDDQNSIK 407

RESULT 6
T15348
hypothetical protein B0350.1 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
C;Accession: T15348
R;Gattung, S.
submitted to the EMBL Data Library, February 1996
A;Description: The sequence of C. elegans cosmid B0350.
A;Reference number: Z19732
A;Accession: T15348
A;Status: preliminary; translated from GB/EMBL/D08J
A;Molecule type: DNA
A;Residues: 1-5170 <GAI>
A;Cross-references: EMBL:U51971; NID:g1208971; PID:g1208972; PIDN:AAA...; GSPDB:...
C;Genetics:
A;Gene: CESP:B0350.1
A;Introns: 48/1; ...; ...; ...;

Query Match 78.3%; Score 36; DB 2; Length ...;
Best Local Similarity 87.5%; Prod. No. 2.3e+02;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0

Qy 2 DDQESVKS 9
|||| :||
Db 2114 DDQESLKS 2126

RESULT 7
A45691
hypothetical protein [imported] - Lactobacillus delbrueckii subsp. [illegible]
C;Species: Lactobacillus delbrueckii subsp. lactis phage LL-H
C;Date: 24-Feb-1994 #sequence_revision 25-Apr-1997 #text_change 28-...-...
C;Accession: A45691
R;Vasala, A.; Dupont, L.; Baumann, M.; Kitzenthaler, P.; Alatossava, ...
J. Virol. 67, 3061-3068, 1993

A;Title: Molecular comparison of the structural proteins encoding gene clusters of two r
A;Reference number: A45691; MUID:93267750
A;Accession: A45691
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-124 <VAS>
A;Cross-references: GB:L4960; NID:g151696; PIDN:AAC05058.1; PID:g623086
A;Note: sequence extracted from NCBI backbone (NCBIN.13663, NCBIF.132665)

Query Match 76.1%; Score 35; DB 2; Length 124;
Best Local Similarity 59.6%; Pred. No. 7.9;
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YDDxESVKS 9
|||:|:|:|
Db 46 YDEQEAIKN 54

RESULT 8
C95115
cytochrome c type biogenesis protein CcdA [imported] - Streptococcus pneumoniae (strain
C;Species: Streptococcus pneumoniae
C;Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 24-Aug-2001
C;Accession: C95115
R;Tettelin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Hold
on, J.D.; Umayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapple,
nson, J.; Hickey, E.K.; Bult, C.E.
Science 293, 498-506, 2001
A;Authors: Tettelin, H.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison,
A;Title: Complete Genome Sequence of a virulent isolate of Streptococcus pneumoniae.
A;Reference number: A96006; MUID:21357209; PMID:11146616
A;Accession: C95115
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-235 <KTS>
A;Cross-references: GB:AE005672; GB:AAK75110.1; PID:g14972412; GSPDB:GN00164; InterPro
A;Experimental source: strain TIGR4
C;Genetics:
A;Gene: SP0999
C;Superfamily: Haemophilus influenzae probable holocytochrome c synthase

Query Match 76.1%; Score 35; DB 2; Length 235;
Best Local Similarity 87.5%; Pred. No. 16;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 DDQESVKS 9
|||||:||
Db 46 DDQESAKS 53

RESULT 9
G97984
cytochrome c type biogenesis protein [imported] - Streptococcus pneumoniae (strain R6)
C;Species: Streptococcus pneumoniae
C;Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 02-Nov-2001
C;Accession: G97984
R;Hoskins, J.; Alborn, W.E., Jr.; Arnold, J.; Blaszczak, L.C.; Burgett, S.; DeHoff, B.S
e, R.; LeBlanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S.; M
y, P.; Sun, P.M.; Winkler, M.E.
J. Bacteriol. 183, 5709-5717, 2001
A;Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Bobb, K.L.; Jaskunas, S.R.;
A;Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.
A;Reference number: A97872; MUID:21429245; PMID:11544234
A;Accession: G97984
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-235 <KHS>
A;Cross-references: GB:AE007317; PIDN:AAK99707.1; PID:g15458502; GSPDB:GN00174
C;Genetics:
A;Gene: ccdA
C;Superfamily: Haemophilus influenzae probable holocytochrome c synthase

Query Match 76.1%; Score 35; DB 2; Length 235;
Best Local Similarity 87.5%; Pred. No. 16;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 DDQESVKS 9
|||||:||
Db 46 DDQESAKS 53

RESULT 10
S61551
breast-regressing protein brp39 precursor - mouse
C;Species: Mus musculus (house mouse)
C;Date: 19-Mar-1997 #sequence_revision 25-Apr-1997 #text_change 26-Jun-200
C;Accession: S61551; S61550; I48271
R;Morrison, B.W.; Leder, P.
Oncogene 9, 3417-3426, 1994
A;Title: neu and ras initiate murine mammary tumors that share genetic markers genera
A;Reference number: I48271; MUID:95060797
A;Accession: S61551
A;Status: nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1-405 <MOR1>
A;Cross-references: EMBL:X93035; NID:g1085065; PIDN:CAA63603.1; PID:g1085066
R;Morrison, B.W.
submitted to the EMBL Data Library, November 1995
A;Reference number: S61550
A;Accession: S61550
A;Molecule type: mRNA
A;Residues: 1-245,'I',247-330,'H',332-350,'MVWALLLLDFQGTCVFKEFFFLINAIKDALA' <MOR2>
A;Cross-references: EMBL:X93035; NID:g1085065; PIDN:CAA63603.1; PID:g1085066
A;Note: the differences at the carboxyl end are due to a frameshift error
C;Genetics:
A;Gene: brp39
C;Superfamily: Streptomyces chitinase chi40
F;1-21/Domain: signal sequence #status predicted <SIG>
F;22-405/Product: breast regressing protein brp39 #status predicted <MAT>

Query Match 76.1%; Score 35; DB 2; Length 405;
Best Local Similarity 66.7%; Pred. No. 29;
Matches 6; Conservative 3; Mismatches 0; Indels 2; Gaps 2;

Qy 1 YDDxESVKS 9
|:|:|:|:|
Db 329 YEDKESVKN 337

RESULT 11
S60618
4-alpha-glucanotransferase (EC 2.4.1.25) - Thermotoga maritima (strain MSB8)
C;Species: Thermotoga maritima
C;Date: 27-Apr-1996 #sequence_revision 08-Nov-1996 #text_change 21-Jul-2000
C;Accession: S60618; S23198; D72487
R;Beinrich, P.; Huber, W.; Liebl, W.
Syst. Appl. Microbiol. 17, 297-305, 1994
A;Title: Expression in Escherichia coli and structure of the gene encoding 4-alpha-gl
related enzyme subfamilies.
A;Reference number: S60618
A;Accession: S60618
A;Molecule type: DNA
A;Residues: 1-441 <HEI>
A;Cross-references: EMBL:Z50813; NID:g951310; PIDN:CAA90693.1; PID:g951311
R;Liebl, W.; Feil, R.; Gabelsberger, J.; Kellermann, J.; Schleifer, K.H.
Eur. J. Biochem. 207, 81-88, 1992
A;Title: Purification and characterization of a novel thermostable 4-alpha-glucanotra
A;Reference number: S23198; MUID:92331687
A;Accession: S23198
A;Molecule type: protein
A;Residues: 'A'-2-23 <LIE>
R;Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hic

Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.;
C.M.
Nature 399, 323-329, 1999
A;Title: Evidence for lateral gene transfer between Archaea and bacteria from genome seq
A;Reference number: A72200; MUID:99287816
A;Accession: D72307
A;Molecule type: DNA
A;Residues: 1-264,'G',266-441 <ARN>
A;Cross-references: PIR:AF0710; GB:AF002512; NID:g1256873; PIDN:AAB95451.1; PID:g498096
A;Experimental source: strain MSB8
C;Genetics:
A;Gene: TM0364
C;Keywords: glycosyltransferase; hexosyltransferase

Query Match 73.9%; Score 34; DB 2; Length 441;
Best Local Similarity 75.0%; Pred. No. 50;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 YDDQESVK 8
|||| |:|
Db 409 YDDQHSLK 416

RESULT 12
S59969
DNA topoisomerase (ATP-hydrolyzing) (EC 5.99.1.3) isoform beta - Chinese hamster
N;Alternate names: DNA topoisomerase II isoform beta; DNA-gyrase
C;Species: Cricetulus griseus (Chinese hamster)
C;Date: 15-Feb 1996 #sequence_revision 13-Mar-1997 #text_change 16-Jul 1999
C;Accession: S59969; S54154
R;Dereuddre, S.; Frey, S.; Delaporte, C.; Jacquemin-Sablon, A.
Biochim. Biophys. Acta 1264, 178-182, 1995
A;Title: Cloning and characterization of full-length cDNAs coding for the DNA topoisomer
A;Reference number: S59969; MUID:96085121
A;Accession: S59969
A;Status: translation not shown
A;Molecule type: mRNA
A;Residues: 1-1612 <DER>
A;Cross-references: EMBL:X86455; NID:g790987; PIDN:CAA60173.1; PID:g790988
A;Experimental source: lung
C;Superfamily: eukaryotic type II DNA topoisomerase; phage T4 DNA topoisomerase (ATP hyd
C;Keywords: ATP; DNA binding; DNA replication; isomerase; nucleus
F;697-927/Domain: phage T4 DNA topoisomerase (ATP hydrolyzing) medium chain homology <T4

Query Match 73.9%; Score 34; DB 2; Length 1612;
Best Local Similarity 66.7%; Pred. No. 2.1e-02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 YDDQESVKS 9
:|| ||:|:
Db 541 YDDAESLKI 549

RESULT 13
A39242
DNA topoisomerase (ATP hydrolyzing) (EC 5.99.1.3) beta, splice form 2 - human
N;Alternate names: DNA topoisomerase II isoform beta 2
N;Contains: DNA topoisomerase II isoform beta-1
C;Species: Homo sapiens (man)
C;Date: 04-Oct-1991 #sequence_revision 03-May-1996 #text_change 21-Jul 2000
C;Accession: S26730; A39242; S10710; S33970; S40190; S41641; S41640
R;Jenkins, J.R.; Ayton, P.; Jones, T.; Davies, S.L.; Simmons, D.L.; Harris, A.L.; Sheer,
Nucleic Acids Res. 20, 5587-5592, 1992
A;Title: Isolation of cDNA clones encoding the beta isozyme of human DNA topoisomerase
A;Reference number: S26730; MUID:93087165
A;Accession: S26730
A;Molecule type: mRNA
A;Residues: 1-23,29-1626 <JEN>
A;Cross-references: EMBL:X68060; NID:g37230; PIDN:CAA48197.1; PID:g7231
R;Chung, T.D.Y.; Drake, F.H.; Tan, K.B.; Per, S.R.; Crooke, S.T.; Mirabelli, C.K.
Proc. Natl. Acad. Sci. U.S.A. 86, 9431-9435, 1989
A;Title: Characterization and immunological identification of cDNA c...Lec...encoding ...
A;Reference number: A39242; MUID:90008281
A;Accession: A39242
A;Molecule type: mRNA
A;Residues: 149-1043 <DHI>
A;Cross-references: GB:M27504
R;Austin, C.A.; Fisher, L.M.
FEBS Lett. 266, 115-117, 1990
A;Title: Isolation and characterization of a human cDNA clone encoding a novel DNA ...
A;Reference number: S10710; MUID:90306433
A;Accession: S10710
A;Molecule type: mRNA
A;Residues: 1043-1276 <AUS>
A;Cross-references: GB:X53662; GB:S56813; NID:g48324; PIDN:CAA37...; PID:g48...
R;Austin, C.A.; Sng, J.H.; Patel, S.; Fisher, L.M.
Biochim. Biophys. Acta 1172, 283-291, 1993
A;Title: Novel HeLa topoisomerase II is the II beta isoform: complete ...d...l sequence
A;Reference number: S40190; MUID:93192319
A;Accession: S33970
A;Molecule type: mRNA
A;Residues: 1-23,29-1610,'A',1612-1626 <AUS>
A;Cross-references: EMBL:Z15111
R;Austin, C.A.; Sng, J.H.; Patel, S.; Fisher, L.M.
submitted to the EMBL Data Library, September 1992
A;Reference number: S40191
A;Accession: S30191
A;Molecule type: mRNA
A;Residues: 596-1430,'S',1432-1610,'A',1612-1626 <AUS>
A;Cross-references: EMBL:Z15110; NID:g288564; PIDN:CAA78821.1; PID:g288565
R;Davies, S.L.; Jenkins, J.R., Hickson, I.D.
Nucleic Acids Res. 21, 3719-3724, 1993
A;Title: Human cells express two differentially spliced forms of topoisomerase II bet
A;Reference number: S41641; MUID:93376494
A;Accession: S41641
A;Molecule type: DNA
A;Residues: 24-80 <DAV>
A;Cross-references: EMBL:X71911; NID:g396540; PIDN:CAA50726.1; PID:g396...
A;Note: this sequence represents a low mitosis splice form, designated beta-
C;Genetics:
A;Gene: GDB:TOP2B
A;Cross-references: GDB:131575; OMIM:126431
A;Map position: 3p24-3p24
C;Superfamily: eukaryotic type II DNA topoisomerase; phage T4 DNA ...; ...merase (ATP
C;Keywords: alternative splicing; ATP; dimer; isomerase; nucleus
F;1-1626/Product: DNA topoisomerase II beta 2 #status predicted <MIN>
F;1-23,29-1626/Product: DNA topoisomerase II beta 1 #status predicted <MAT>
F;709-949/Domain: phage T4 DNA topoisomerase (ATP hydrolyzing) medium chain homology

Query Match 73.9%; Score 34; DB 2; Length 1626;
Best Local Similarity 66.7%; Pred. No. 2.1e-02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 YDDQESVKS 9
:|| ||:|:
Db 543 YDDAESLKI 551

RESULT 14
C81674
hemolysin, probable TC0706 [imported] - Chlamydia muridarum (strain Nigg)
C;Species: Chlamydia muridarum, Chlamydia trachomatis Mouse
C;Date: 31-Mar-2000 #sequence_revision 01-Mar-2000 #text_change 21-Jul-...
C;Accession: C81674
R;Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, ...
, C.; Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, ...
Nucleic Acids Res. 28, 1397-1406, 2000
A;Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR...
A;Reference number: A81500; MUID:20150255
A;Accession: C81674
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-374 <PET>

A;Cross reference: GB:AE000439; GB:AE032160; NID:g71907; PIDN:AAF59520.1; PID:g71907
A;Experimental source: strain Nissle (MoPn)
C;Genetics:
A;Gene: ECOTOX

Query Match 71.7%; Score 83; DB 2; Length 574;
Best Local Similarity 85.7%; Pred. No. 67;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 YDDQESV 7
|||||| |
Db 277 YDDQEDV 283

FESOLI 15
F64845
paraquat-inducible protein pqiB - Escherichia coli
C;Species: Escherichia coli
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 21-Jul-2000
C;Accession: F64845; B57146; S52912
R;Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co
.A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A;Title: The complete genome sequence of Escherichia coli K-12.
A;Reference number: A64220; MUID:97426617
A;Accession: F64845
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-540 <BLA>
A;Cross-references: GB:AE000196; GB:U00096; NID:g1787180; PIDN:AAC74017.1; PID:g1787184;
A;Experimental source: strain K-12, substrain MG1655
R;Koh, Y.S.; Roe, J.-H.
J. Bacteriol. 177, 2674-2678, 1995
A;Title: Isolation of a novel paraquat-inducible (pqi) gene regulated by the soxRS locus
A;Reference number: A57146; MUID:95273582
A;Accession: B57146
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-791 <KOH>
A;Cross-references: EMBL:X81561
A;Experimental source: strain K-12, substrain W3110
C;Comment: Residues 1-282 appear to be homologous to PIR:B64945 and other members of tha
C;Genetics:
A;Gene: pqiB; pqi5B
C;Superfamily: pqiB protein
C;Keywords: transmembrane protein
F;19-35/Domain: transmembrane #status predicted <TMM>

Query Match 71.7%; Score 83; DB 1; Length 546;
Best Local Similarity 62.5%; Pred. No. 1e-02;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YDDQESVK 8
|:|||:::
Db 271 YDDGKSIQ 278

Search completed: August 26, 2002, 06:11:42
Job time: 108 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 26, 2002, 06:29:02 ; Search time 35.99 Seconds
(without alignments)
9.683 Million cell updates/sec

Title: US-09-744-282-3
Perfect score: 46
Sequence: 1 YDDQESVKS 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	46	100.0	383	1	C3L1_HUMAN	P36222 homo sapien
2	46	78.3	797	1	DPOM_AGABT	P30322 agaricus bi
3	45	76.1	124	1	VSP1_BPLLH	Q04765 lactococcus
4	45	76.1	1553	1	IP2A_CHICK	Q42140 gallus gall
5	44	73.9	441	1	MG1A_THEMA	P80099 thermotoga
6	44	73.9	955	1	VP2_BTV17	P05309 bluetongue
7	44	73.9	1612	1	IP2B_CRILO	Q64399 cricetulus
8	44	73.9	1612	1	IP2B_MOUSE	Q64511 mus musculu
9	44	73.9	1626	1	IP2B_HUMAN	Q02980 homo sapien
10	43	71.7	546	1	PQIB_ECOLI	P43671 escherichia
11	43	71.7	827	1	RAP1_YEAST	P11938 saccharomyc
12	43	71.7	1627	1	IP2B_CHICK	Q42131 gallus gall
13	42	69.6	128	1	VSP1_BPMV4	Q04766 lactococcus
14	42	69.6	339	1	RGG_CAMJE	Q9pki0 campylobact
15	42	69.6	494	1	AMY1_SACFI	P21567 saccharomyc
16	42	69.6	518	1	YX23_CAEEL	Q11194 caenorhabdi
17	42	69.6	655	1	GR78_PLAFC	Q05866 plasmodium
18	42	69.6	773	1	MAK5_YEAST	P38712 saccharomyc
19	42	69.6	875	1	HELX_SULSO	P95949 sulfolobus
20	42	69.6	956	1	VP2_BTV11	P05308 bluetongue
21	41	67.4	256	1	DP3E_HAEIN	P43745 haemophilus
22	41	67.4	329	1	F16Q_BETVU	Q42649 beta vulgar
23	41	67.4	331	1	F16Q_SACHY	Q43139 saccharum h
24	41	67.4	339	1	F16Q_BRANA	P46267 brassica na
25	41	67.4	341	1	F16Q_SPIOL	P14766 spinacia ol
26	41	67.4	377	1	LMXR_CHICK	P53413 gallus gall
27	41	67.4	448	1	BIOA_BACSU	P53555 bacillus su
28	41	67.4	499	1	CHIB_SERMA	P11797 serratia ma
29	41	67.4	699	1	CHI1_BACCI	P20533 bacillus ci
30	41	67.4	777	1	PAR1_HUMAN	Q99728 homo sapien
31	41	67.4	1211	1	SSC2_PICAN	Q96ab6 pichia ang
32	41	67.4	1650	1	BP28_CAEEL	Q23495 caenorhabdi
33	41	67.4	2103	1	KRPL_DUCK	P23403 anas platyr
34	30	65.2	98	1	RL21_HALN1	Q9hqq8 halobacteri
35	30	65.2	116	1	LUXE_PHOLU	P19842 photorhabdu
36	30	65.2	173	1	IPYK_SULAC	P50388 sulfolobus
37	30	65.2	244	1	DCG1_YEAST	P32460 saccharomyc
38	30	65.2	267	1	RL4_SULSO	Q9ux86 sulfolobus
39	30	65.2	312	1	CAB4_BOVIN	Q95323 bos taurus
40	30	65.2	340	1	F16Q_SOLTU	P46276 solanum tub
41	30	65.2	346	1	MPA1_CUPAR	Q9sed9 cupressus a
42	30	65.2	367	1	MPA1_JUNAS	P81294 juniperus a
43	30	65.2	411	1	CBF1_YARLI	P45815 yarrowia li
44	30	65.2	430	1	GSA_BACSU	P30949 bacillus su
45	30	65.2	447	1	MURE_HELPY	O26027 helicobacte

ALIGNMENTS

RESULT 1
C3L1_HUMAN
ID C3L1_HUMAN STANDARD; PRT; 383 AA.
AC P36222; P30923;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Chitinase-3 like protein 1 precursor (Cartilage glycoprotein-39) (GP-
DE 39) (39 kDa synovial protein) (YKL-40).
GN CHI3L1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC TISSUE=Cartilage;
RX MEDLINE 94064659; PubMed=8215017;
RA Hakala B.E., White C., Recklies A.D.;
RT "Human cartilage gp-39, a major secretory product of articular
RT chondrocytes and synovial cells, is a mammalian member of a chitinase
RT protein family.";
RL J. Biol. Chem. 268:25803-25810(1993).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Blood;
RX MEDLINE=97296591; PubMed=9244440;
RA Rehli M., Krause S.W., Andreesen R.;
RT "Molecular characterization of the gene for human cartilage gp-39
RT (CHI3L1), a member of the chitinase protein family and marker for
RT late stages of macrophage differentiation.";
RL Genomics 43:221-225(1997).
RN [3]
RP SEQUENCE OF 22-45.
RX MEDLINE 90328983; PubMed=2375755;
RA Nyirkos P., Golds E.E.;
RT "Human synovial cells secrete a 39 kDa protein similar to a bovine
RT mammary protein expressed during the non-lactating period.";
RL Biochem. J. 269:265-268(1990).
CC -!- FUNCTION: MAY PLAY AN IMPORTANT ROLE IN THE CAPACITY OF CELLS TO
CC RESPOND TO AND COPE WITH CHANGES IN THEIR ENVIRONMENT.
CC -!- SUBCELLULAR LOCATION: Extracellular.
CC -!- TISSUE SPECIFICITY: PRESENT IN ARTICULAR CHONDROCYTES, SYNOVIAL
CC CELLS AS WELL AS IN LIVER. UNDETECTABLE IN MUSCLE TISSUES, LUNG,
CC PANCREAS, MONONUCLEAR CELLS, OR FIBROBLASTS.
CC -!- PTM: GLYCOSYLATED.
CC -!- SIMILARITY: BELONGS TO FAMILY 18 OF GLYCOSYL HYDROLASES.
CC ----------
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC ----------

```
DR   EMBL; M80927; AAA16074.1; -.
DR   EMBL; Y08374; CAA69661.1; -.
DR   EMBL; Y08375; CAA69661.1; JOINED.
DR   EMBL; Y08376; CAA69661.1; JOINED.
DR   EMBL; Y08377; CAA69661.1; JOINED.
DR   EMBL; Y08378; CAA69661.1; JOINED.
DR   PIR; S10677; S10677.
DR   PIR; A43162; A43162.
DR   MIM; 601525; -.
DR   InterPro; IPR001579; Chitinase_2.
DR   InterPro; IPR001223; Glyco_hydro_18.
DR   Pfam; PF00704; Glyco_hydro_18; 1.
DR   PROSITE; PS01095; CHITINASE_18; FALSE_NEG.
KW   Glycoprotein; Signal.
FT   SIGNAL        1     21
FT   CHAIN        22    383       CHITINASE-3 LIKE PROTEIN 1.
FT   CARBOHYD     60     60       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   383 AA;  42613 MW;  76ADD8298EEEC2D1 CRC64;


  Query Match             100.0%;  Score 46;  DB 1;  Length 383;
  Best Local Similarity   100.0%;  Pred. No. 0.06;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 YDDQESVKS 9
           |||||||||
Db     328 YDDQESVKS 336


RESULT   2
DPOM_AGABT
ID   DPOM_AGABT     STANDARD;      PRT;   797 AA.
AC   P30322;
DT   01-APR-1993 (Rel. 25, Created)
DT   01-APR-1993 (Rel. 25, Last sequence update)
DT   15-DEC-1998 (Rel. 37, Last annotation update)
DE   Probable DNA polymerase (EC 2.7.7.7) (Fragment).
OS   Agaricus bitorquis.
OG   Mitochondrion.
OG   Plasmid pEM.
OC   Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Homobasidiomycetes.
OC   Agaricales; Agaricaceae; Agaricus.
OX   NCBI_TaxID=5343;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN ATCC 24666 / AG4;
RX   MEDLINE=92147410; PubMed=1979091;
RA   Robison M.M., Royer J.C., Horgen P.A.;
RT   "Homology between mitochondrial DNA of Agaricus bisporus and an
RT   internal portion of a linear mitochondrial plasmid of Agaricus
RT   bitorquis.";
RL   Curr. Genet. 19:495-502(1991).
CC   -!- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate
CC       + {DNA}(N).
CC   -!- MISCELLANEOUS: THIS DNA POLYMERASE REQUIRES A PROTEIN AS A PRIMER
CC       (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO DNA POLYMERASE TYPE B FAMILY. STRONG, TO
CC       DNA POLYMERASE OF OTHER FUNGAL AND PLANTS MITOCHONDRIAL PLASMIDS.
CC   ---------- ------ ------ ----------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------- ------ ------ ----------------------------------------------
DR   EMBL; X63075; CAA44800.1; -.
DR   PIR; S28103; S28103.
DR   InterPro; IPR002064; DNA_pol_B.
DR   SMART; SM00486; POLBc; 1.
DR   PROSITE; PS00116; DNA_POLYMERASE_B; FALSE_NEG.
KW   Transferase; DNA directed DNA polymerase; DNA replication;
KW   DNA-binding; Plasmid; Mitochondrion.
FT   NON_TER       1       1
SQ   SEQUENCE   797 AA;  91922 MW;  9C40EAF51FB7F927 CRC64;


  Query Match             78.3%;  Score 36;  DB 1;  Length 797;
  Best Local Similarity   75.0%;  Pred. No. 14;
  Matches    6;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 YDDQESVK 8
           |||| |:|
Db     400 YDDQNSIK 407


RESULT   3
VSP1_BPLLB
ID   VSP1_BPLLB     STANDARD;      PRT;   124 AA.
AC   Q04765;
DT   01-OCT-1993 (Rel. 27, Created)
DT   01-OCT-1993 (Rel. 27, Last sequence update)
DT   01-FEB-1994 (Rel. 28, Last annotation update)
DE   Structural protein (ORF1) (Fragment).
OS   Lactococcus delbrueckii bacteriophage LL-H.
OC   Viruses.
OX   NCBI_TaxID=12348;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=93267750; PubMed=8497043;
RA   Vasala A., Dupont L., Baumann M., Ritzenthaler P., Alatossava T.;
RT   "Molecular comparison of the structural proteins encoding gene
RT   clusters of two related Lactobacillus delbrueckii bacteriophages ";
RL   J. Virol. 67:3061-3068(1993).
CC   -!- SIMILARITY: TO BACTERIOPHAGE MV4 STRUCTURAL PROTEIN.
CC   ---------- ------ ------ ----------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------- ------ ------ ----------------------------------------------
DR   EMBL; L02496; AAC00558.1; -.
DR   PIR; A45691; A45691.
KW   Structural protein.
FT   NON_TER       1       1
SQ   SEQUENCE   124 AA;  14727 MW;  9A1A6C0AC952123 CRC64;


  Query Match             76.1%;  Score 35;  DB 1;  Length 124;
  Best Local Similarity   55.6%;  Pred. No. 2.8;
  Matches    5;  Conservative    4;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 YDDQESVKS 9
           |:|:|::|:
Db     46 YDEQEAIKN 54


RESULT   4
TP2A_CHICK
ID   TP2A_CHICK     STANDARD;      PRT;  1553 AA.
AC   O42130;
DT   15-JUL-1998 (Rel. 36, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   DNA topoisomerase II, alpha isozyme (EC 5.99.1.3).
GN   TOP2A.
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=9031;
RN   [1]
```

```
RP   SEQUENCE FROM N.A.
RA   Nitmi A., Brrat M., Mitsui J.;
RT   "Chicken DNA topoisomerase 2 alpha and beta.";
RL   Submitted (JAN-2000) to the EMBL/GenBank/DDBJ databases.
CC   !- FUNCTION: CONTROLS THE TOPOLOGICAL STATES OF DNA BY TRANSIENT
CC       BREAKAGE AND SUBSEQUENT REJOINING OF DNA STRANDS. TOPOISOMERASE II
CC       MAKES DOUBLE STRAND BREAKS (BY SIMILARITY).
CC   !- CATALYTIC ACTIVITY: ATP-dependent breakage, passage and rejoining
CC       of double-stranded DNA.
CC   -!- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: NUCLEAR; GENERALLY LOCATED IN THE
CC       NUCLEOPLASM.
CC   -!- MISCELLANEOUS: EUKARYOTIC TOPOISOMERASE I AND II CAN RELAX BOTH
CC       NEGATIVE AND POSITIVE SUPER COILS. WHEREAS TOPOISOMERASE I CAN
CC       RELAX ONLY NEGATIVE SUPERCOILS.
CC   !   SIMILARITY: BELONGS TO THE TYPE II TOPOISOMERASE FAMILY.
CC   ----------- -- -- ------- --- ---- -------- ----- -- - ----------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --- ---------- ------- --- --------- ---- ---- ------ --
DR   EMBL; AB007445; BAA22539.2; -.
DR   HSSP; P06786; 1PGW.
DR   InterPro; IPR003997; CBFA_NFYB_topis.
DR   InterPro; IPR001241; DNA_topoisoII.
DR   InterPro; IPR002205; DNA_topoisoIV.
DR   InterPro; IPR002594; HATPase_c.
DR   Pfam; PF00204; DNA_topoisoII; 1.
DR   Pfam; PF00521; DNA_topoisoIV; 1.
DR   Pfam; PF02518; HATPase_c; 1.
DR   PRINTS; PR00418; TPI2FAMILY.
DR   ProDom; PD000616; DNA_topoisoII; 1.
DR   SMART; SM00387; HATPase_c; 1.
DR   SMART; SM00433; TOP2c; 1.
DR   SMART; SM00434; TOP4c; 1.
DR   PROSITE; PS00177; TOPOISOMERASE_II; 1.
KW   Isomerase; Topoisomerase; DNA binding; ATP binding; Nuclear protein.
FT   NP BIND     26     167       ATP (POTENTIAL).
FT   ACT_SITE    805    805       DNA CLEAVAGE (BY SIMILARITY).
SQ   SEQUENCE   1554 AA;  174997 MW;  3322CE95238F71B2 CRC64;


  Query Match             75.1%;  Score 35;  DB 1;  Length 1554;
  Best Local Similarity   66.7%;  Pred. No. 48;
  Matches    6;  Conservative    4;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 YDDQESVKS 9
           |:|:||:|
Db     523 YEDKESLKS 531


RESULT   5
MGIA_THEMA
ID   MGIA_THEMA      STANDARD;      PRT:   441 AA.
AC   P05309;
DT   01-AUG-1992 (Rel. 17, Created)
DT   30-MAY-2000 (Rel. 38, Last sequence update)
DT   15-OCT-2001 (Rel. 40, Last annotation update)
DE   4-alpha-glucanotransferase (EC 2.4.1.25) (Amylomaltase)
DE   (Disproportionating enzyme) (D enzyme).
GN   MGIA OR TM0364.
OS   Thermotoga maritima.
OC   Bacteria; Thermotogales; Thermotogaceae.
OX   NCBI_TaxID=2336;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MSB8 / DSM 3109;
RX   MEDLINE=99287577; PubMed=10360571;
```

```
RA   Nelson K.E., Clayton R.A., Gill S.R., Gwinn M.L., Dodson R.J.,
RA   Haft D.H., Hickey E.K., Peterson J.D., Nelson W.C., Ketchum K.A.,
RA   McDonald L., Utterback T.R., Malek J.A., Linher K.D., Garrett M.M.,
RA   Stewart A.M., Cotton M.D., Pratt M.S., Phillips C.A., Richardson D.,
RA   Heidelberg J., Sutton G.G., Fleischmann R.D., Eisen J.A., White O.,
RA   Salzberg S.L., Smith H.O., Venter J.C., Fraser C.M.;
RT   "Evidence for lateral gene transfer between Archaea and Bacteria from
RT   genome sequence of Thermotoga maritima.";
RL   Nature 399:323-329(1999).
RN   [2]
RP   SEQUENCE OF 1-23, AND CHARACTERIZATION.
RC   STRAIN=MSB8 / DSM 3109;
RX   MEDLINE=92321687; PubMed=1628064;
RA   Jolff W., Hilt K., Sahleinstein ..., Kellermann J., Schleifer K.-H.;
RT   "Purification and characterization of a novel thermostable 4-alpha-
RT   glucanotransferase of Thermotoga maritima cloned in Escherichia
RT   coli.";
RL   Eur. J. Biochem. 207:81-88(1992).
CC   !- CATALYTIC ACTIVITY: Transfers a segment of a (1,4)-alpha D-glucan
CC       to a new 4-position in an acceptor, which may be glucose or (1,4)-
CC       alpha D-glucan.
CC   !- SUBUNIT: MONOMER.
CC   !   SUBCELLULAR LOCATION: Cytoplasmic.
CC   -!- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
CC       KNOWN AS THE ALPHA-AMYLASE FAMILY.
CC   ----------- -- --- ------- --------------------- ----- -- - ----------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --- ---------- ------- --- --------- ---- ---- ------ --
DR   EMBL; AE001716; AAD35451.1; -.
DR   PIR; S23198; S23198.
DR   TIGR; TM0364; -.
DR   InterPro; IPR000461; Alpha_amylase.
DR   Pfam; PF00128; alpha-amylase; 1.
DR   PRINTS; PR00110; ALPHAAMYLASE.
KW   Transferase; Glycosyltransferase; Carbohydrate metabolism;
KW   Complete proteome.
FT   ACT_SITE    216    216       BY SIMILARITY.
FT   ACT_SITE    278    278       BY SIMILARITY.
FT   CONFLICT      1      1       M -> A (IN REF. 2).
SQ   SEQUENCE   441 AA;  51843 MW;  9FB4C2ABC9D3DF3A CRC64;


  Query Match             74.9%;  Score 34;  DB 1;  Length 441;
  Best Local Similarity   75.0%;  Pred. No. 19;
  Matches    6;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 YDDQESVK 8
           |||| |:|
Db     409 YDDQHSLK 416


RESULT   6
VP2_BTV17
ID   VP2_BTV17       STANDARD;      PRT:   955 AA.
AC   P05309;
DT   01-NOV-1988 (Rel. 09, Created)
DT   01-NOV-1988 (Rel. 09, Last sequence update)
DT   15-DEC-1998 (Rel. 37, Last annotation update)
DE   Outer capsid protein VP2.
GN   S2.
OS   Bluetongue virus (serotype 17 / isolate USA).
OC   Viruses; dsRNA viruses; Reoviridae; Orbivirus.
OX   NCBI_TaxID=33718;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=87212094; PubMed=2820123;
RA   Ghiasi H., Fukusho A., Eshita Y., Roy P.;
```

```
RT   "Identification and characterization of conserved and variable
RT   regions in the neutralization VP2 gene of bluetongue virus.";
RL   Virology 160:100-109(1987).
CC   -!- FUNCTION: THE VP2 PROTEIN IS ONE OF THE TWO PROTEINS (WITH VP5)
CC       WHICH CONSTITUTE THE VIRUS PARTICLE OUTER CAPSID. IT IS THE
CC       MAJOR TARGET OF THE HOST IMMUNOGENIC RESPONSE.
CC   -!- SIMILARITY: BELONGS TO THE REOVIRUSES VP2 FAMILY.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; M17438; AAA42830.1; -.
DR   InterPro; IPR001742; Orbi_VP2.
DR   Pfam; PF00898; Orbi_VP2; 1.
DR   ProDom; PD002948; Orbi_VP2; 1.
KW   Coat protein.
SQ   SEQUENCE   955 AA;  110572 MW;  E26CF75862FD6BAA CRC64;


  Query Match             72.9%;  Score 44;  DB 1;  Length 955;
  Best Local Similarity   66.7%;  Pred. No. 44;
  Matches    6;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 YDDQESVKS 9
           ||:::|.||
Db      84 YDDRKSTKS 92


RESULT   7
TP2B_CRILO
ID   TP2B_CRILO     STANDARD;      PRT;  1612 AA.
AC   Q64399;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   DNA topoisomerase II, beta isozyme (EC 5.99.1.3).
GN   TOP2B.
OS   Cricetulus longicaudatus (long-tailed hamster) (Chinese hamster).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC   Cricetulus.
OX   NCBI_TaxID=10030;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Lung;
RX   MEDLINE=96095121; PubMed=7495861;
RA   Deroanddre S., Frey S., Delaporte C., Jacquemin-Sablon A.;
RT   "Cloning and characterization of full-length cDNAs coding for the DNA
RT   topoisomerase II beta from Chinese hamster lung cells sensitive and
RT   resistant to 9-OH-ellipticine.";
RL   Biochim. Biophys. Acta 1264:178-182(1995).
CC   -!- FUNCTION: CONTROL OF TOPOLOGICAL STATES OF DNA BY TRANSIENT
CC       BREAKAGE AND SUBSEQUENT REJOINING OF DNA STRANDS. TOPOISOMERASE II
CC       MAKES DOUBLE-STRAND BREAKS.
CC   -!- CATALYTIC ACTIVITY: ATP-dependent breakage, passage and rejoining
CC       of double-stranded DNA.
CC   -!- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: NUCLEAR; NUCLEOLUS.
CC   -!- MISCELLANEOUS: EUKARYOTIC TOPOISOMERASE I AND II CAN RELAX BOTH
CC       NEGATIVE AND POSITIVE SUPERCOILS, WHEREAS PROKARYOTIC ENZYMES
CC       RELAX ONLY NEGATIVE SUPERCOILS.
CC   -!- SIMILARITY: BELONGS TO THE TYPE II TOPOISOMERASE FAMILY.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
```

```
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; X86455; CAA60178.1; -.
DR   HSSP; P06786; 1BGW.
DR   InterPro; IPR000357; CBFA_NFYB_hapls.
DR   InterPro; IPR000221; DNA_topoisoII.
DR   InterPro; IPR000205; DNA_topoisoIV.
DR   InterPro; IPR000594; BATPase_C.
DR   Pfam; PF00204; DNA_topoisoII; 1.
DR   Pfam; PF00521; DNA_topoisoIV; 1.
DR   Pfam; PF02518; BATPase; 1.
DR   PRINTS; PR00615; CCAATSUBUNIA.
DR   PRINTS; PR00119; TOPISMRASE.
DR   ProDom; PD000610; DNA_topoisoII; 1.
DR   SMART; SM00433; TOP2c; 1.
DR   SMART; SM00434; TOP4c; 1.
DR   PROSITE; PS00177; TOPOISOMERASE_II; 1.
KW   Isomerase; Topoisomerase; DNA binding; ATP-binding; Nuclear protein.
FT   NP_BIND     170    175       ATP (POTENTIAL).
FT   ACT_SITE    814    814       DNA CLEAVAGE (BY SIMILARITY).
SQ   SEQUENCE   1612 AA;  182074 MW;  C01DGFC40620FC68 CRC64;


  Query Match             73.0%;  Score 44;  DB 1;  Length 1612;
  Best Local Similarity   66.7%;  Prod. No. 80;
  Matches    6;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 YDDQESVKS 9
           ||.|.|:|:
Db     531 YDDAESLKI 539


RESULT   8
TP2B_MOUSE
ID   TP2B_MOUSE     STANDARD;      PRT;  1612 AA.
AC   Q64511;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   DNA topoisomerase II, beta isozyme (EC 5.99.1.3).
GN   TOP2B.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BALB/C; TISSUE=Brain;
RA   Miyake M., Adachi N., Kikuchi A.;
RL   Submitted (SEP-1994) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: CONTROL OF TOPOLOGICAL STATES OF DNA BY TRANSIENT
CC       BREAKAGE AND SUBSEQUENT REJOINING OF DNA STRANDS. TOPOISOMERASE II
CC       MAKES DOUBLE-STRAND BREAKS.
CC   -!- CATALYTIC ACTIVITY: ATP-dependent breakage, passage and rejoining
CC       of double-stranded DNA.
CC   -!- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: NUCLEAR; NUCLEOLUS.
CC   -!- MISCELLANEOUS: EUKARYOTIC TOPOISOMERASE I AND II CAN RELAX BOTH
CC       NEGATIVE AND POSITIVE SUPERCOILS, WHEREAS PROKARYOTIC ENZYMES
CC       RELAX ONLY NEGATIVE SUPERCOILS.
CC   -!- SIMILARITY: BELONGS TO THE TYPE II TOPOISOMERASE FAMILY.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; D38046; BAA07236.1; -.
DR   HSSP; P06786; 1BGW.
```

```
DR   MEDI; MGI:98791; Top2l.
DR   InterPro; IPR003957; CBFA_NFYB_topis.
DR   InterPro; IPR001241; DNA_topoisoII.
DR   InterPro; IPR002205; DNA_topoisoIV.
DR   InterPro; IPR003594; BATPase_C.
DR   Pfam; PF00204; DNA_topoisoII; 1.
DR   Pfam; PF00521; DNA_topoisoIV; 1.
DR   Pfam; PF02518; HATPase_c; 1.
DR   PRINTS; PR00415; TOPOISMRASE.
DR   PRINTS; PR01416; TOP2FAMILY.
DR   ProDom; PD000016; DNA_topoisoII; 1.
DR   SMART; SM00434; TOP2c; 1.
DR   SMART; SM00434; TOP4c; 1.
DR   PROSITE; PS01177; TOPOISOMERASE_II; 1.
KW   Isomerase; Topoisomerase; DNA binding; ATP binding; Nuclear protein.
FT   NP BIND     170    175       ATP (POTENTIAL).
FT   ACT SITE    814    814       DNA CLEAVAGE (BY SIMILARITY).
SQ   SEQUENCE   1612 AA;  [illegible]

  Query Match          [illegible]%; Score 84; DB 1; Length 1612;
  Best Local Similarit[illegible]%; Pred. No. [illegible];
  Matches    9; Conservat[illegible]%;  2; Mismatches    1; Indels    0; Gaps    0;

Qy       1 YDDQESVKS 9
           ||| | ||:
Db     531 YDDAESLKI 539


RESULT    9
TOP2B_HUMAN
ID   TOP2B_HUMAN     STANDARD;      PRT:  1626 AA.
AC   [illegible]; [illegible]; [illegible]; [illegible];
DT   01-JUL-1993 (Rel. 26, Created)
DT   01-MAR-2002 (Rel. 41, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
DE   DNA topoisomerase II, beta isozyme (EC 5.99.1.3).
GN   TOP2B.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=90087165; PubMed=1333647;
RA   Jenkins J.R., Ayton P., Jones T., Davies S.L., Simmons D.L.,
RA   Harris A.L., Sheer D., Hickson I.D.;
RT   "Isolation of cDNA clones encoding the beta isozyme of human DNA
RT   topoisomerase II and localisation of the gene to chromosome 3p24.";
RL   Nucleic Acids Res. 20:5587-5592(1992).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=94192439; PubMed=8288547;
RA   Austin C.A., Sng J.H., Patel S., Fisher L.M.;
RT   "Novel HeLa topoisomerase II is the II beta isoform: complete coding
RT   sequence and homology with other type II topoisomerases.";
RL   Biochim. Biophys. Acta 1172:283-291(1993).
RN   [3]
RP   SEQUENCE OF 349-1044 FROM N.A.
RX   MEDLINE=90085281; PubMed=2599712;
RA   Chung I.D., Drake F.H., Tan K.B., Per S.R., Crooke S.T.,
RA   Mirabelli C.K.;
RT   "Characterization and immunological identification of cDNA clones
RT   encoding two human DNA topoisomerase II isozymes.";
RL   Proc. Natl. Acad. Sci. U.S.A. 86:9431-9435(1989).
RN   [4]
RP   SEQUENCE OF 910-1626 FROM N.A.
RX   MEDLINE=91196712; PubMed=1009062;
RA   Sng J.H., Heaton V.J., Bell M., Maini P., Austin C.A., Fisher L.M.;
RT   "Molecular cloning and characterization of the human topoisomerase
RT   IIalpha and IIbeta genes: evidence for isoform evolution through gene
RT   duplication.";
RL   Biochim. Biophys. Acta 1444:395-406(1999).

RN   [5]
RP   SEQUENCE OF 1038-1271 FROM N.A.
RX   MEDLINE=90306333; PubMed=2163884;
RA   Austin C.A., Fisher L.M.;
RT   "Isolation and characterization of a human cDNA clone encoding a
RT   novel DNA topoisomerase II homologue from HeLa cells.";
RL   FEBS Lett. 266:115-117(1990).
RN   [6]
RP   SEQUENCE OF 1277-1626 FROM N.A.
RA   Yuwen H.;
RT   "Binding of wild type p53 by topoisomerase II and overexpression of
RT   topoisomerase II in human hepatocellular carcinoma.";
RL   Submitted (APR-1996) to the EMBL/GenBank/DDBJ databases.
RN   [7]
RP   ALTERNATIVE SPLICING.
RX   MEDLINE=93376494; PubMed=8396237;
RA   Wachs F.L., Jenkins J.R., Hickson I.D.;
RT   "Human cells express two differentially spliced forms of
RT   topoisomerase II beta mRNA.";
RL   Nucleic Acids Res. 21:3719-3723(1993).
CC   -!- FUNCTION: CONTROL OF TOPOLOGICAL STATES OF DNA BY TRANSIENT
CC       BREAKAGE AND SUBSEQUENT RESTORATION OF DNA STRANDS. DEFICIENCY IN
CC       MAKES DOUBLE-STRAND BREAKS.
CC   -!- CATALYTIC ACTIVITY: ATP-dependent breakage, passage and rejoining
CC       of double-stranded DNA.
CC   -!- SUBUNIT: HOMODIMER.
CC   -!- SUBCELLULAR LOCATION: NUCLEAR; NUCLEOLUS.
CC   -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS, BETA 1 AND BETA 2 (SHOWN HERE);
CC       ARE PRODUCED BY ALTERNATIVE SPLICING.
CC   -!- PTM: PHOSPHORYLATED.
CC   -!- MISCELLANEOUS: EUKARYOTIC TOPOISOMERASE I AND II CAN RELAX BOTH
CC       NEGATIVE AND POSITIVE SUPERCOILS, WHEREAS PROKARYOTIC ENZYMES
CC       RELAX ONLY NEGATIVE SUPERCOILS.
CC   -!- SIMILARITY: BELONGS TO THE TYPE II TOPOISOMERASE FAMILY.
CC   ---------- [illegible]
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (see http://www.isb-sib.ch/[illegible]
CC   or send an email to license@isb-sib.ch).
CC   ---------- [illegible]
DR   EMBL; X68060; CAA48197.1; -.
DR   EMBL; X71931; [illegible]; NOT_ANNOTATED_CDS.
DR   EMBL; Z15111; CAA78815.1; -.
DR   EMBL; Z15115; CAA78821.1; -.
DR   EMBL; M27504; AAA61210.1; -.
DR   EMBL; AJ011721; CAA09753.1; -.
DR   EMBL; AJ011722; CAA09753.1; JOINED.
DR   EMBL; AJ011723; CAA09753.1; JOINED.
DR   EMBL; AJ011724; CAA09753.1; JOINED.
DR   EMBL; AJ011725; CAA09753.1; JOINED.
DR   EMBL; AJ011726; CAA09753.1; JOINED.
DR   EMBL; AJ011727; CAA09753.1; JOINED.
DR   EMBL; AJ011728; CAA09753.1; JOINED.
DR   EMBL; AJ011729; CAA09753.1; JOINED.
DR   EMBL; AJ011730; CAA09753.1; JOINED.
DR   EMBL; AJ011731; CAA09753.1; JOINED.
DR   EMBL; AJ011732; CAA09753.1; JOINED.
DR   EMBL; X53662; CAA37706.1; -.
DR   EMBL; U54841; AAB01982.1; -.
DR   PIR; S26730; S26730.
DR   PIR; S10710; S10710.
DR   PIR; S41641; S41641.
DR   HSSP; P06786; 1BGW.
DR   MIM; 126431; -.
DR   InterPro; IPR003957; CBFA_NFYB_topis.
DR   InterPro; IPR001241; DNA_topoisoII.
DR   InterPro; IPR002205; DNA_topoisoIV.
DR   InterPro; IPR003594; BATPase_C.
DR   Pfam; PF00204; DNA_topoisoII; 1.
DR   Pfam; PF00521; DNA_topoisoIV; 1.
```

```
DR   Pfam; PF02518; HATPase_c; 1.
DR   PRINTS; PR00615; TOPAISOMRUNTA.
DR   PRINTS; PR00418; TP12FAMILY.
DR   ProDom; PD000616; DNA_topoisoII; 1.
DR   SMART; SM00433; TOP2c; 1.
DR   SMART; SM00434; TOP4c; 1.
DR   PROSITE; PS00177; TOPOISOMERASE_II; 1.
KW   Isomerase; Topoisomerase; DNA-binding; ATP binding; Phosphorylation;
KW   Alternative splicing; Nuclear protein.
FT   NP_BIND     182    187       ATP (POTENTIAL).
FT   ACT_SITE    826    826       DNA CLEAVAGE (BY SIMILARITY).
FT   VARSPLIC     24     28       MISSING (IN ISOFORM BETA 1).
FT   CONFLICT   1441   1441       T -> S (IN REF. 4).
FT   CONFLICT   1611   1611       A -> T (IN REF. 1).
FT   CONFLICT   1621   1621       D -> H (IN REF. 4).
SQ   SEQUENCE   1626 AA;  183266 MW;  E60E9262C68B05D CRC64;


  Query Match             73.9%;  Score 34;  DB 1;  Length 1626;
  Best Local Similarity   66.7%;  Pred. No. 81;
  Matches    6;  Conservative    2;  Mismatches    1;  Indels    1;  Gaps    0;

Qy       1 YDDQESVKS 9
           ||: | :||:
Db     543 YDDAESLKT 551


RESULT  10
PQIB_ECOLI
ID   PQIB_ECOLI     STANDARD;      PRT;   546 AA.
AC   P43671; P77584;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Paraquat inducible protein B.
GN   PQIB OR PQI5B OR B0951.
OS   Escherichia coli.
OC   Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC   Escherichia.
OX   NCBI_TaxID=562;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN K12 / MG1655;
RX   MEDLINE=97426617; PubMed=9278503;
RA   Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA   Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA   Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA   Mau B., Shao Y.;
RT   "The complete genome sequence of Escherichia coli K-12.";
RL   Science 277:1453-1474(1997).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN K1;
RX   MEDLINE=97061202; PubMed=8995232;
RA   Oshima T., Aiba H., Baba T., Fujita K., Hayashi K., Honjo A.,
RA   Ikemoto K., Inada T., Itoh T., Kajihara M., Kanai K., Kashimoto S.,
RA   Kimura S., Kitagawa M., Makino K., Morishita S., Mori H., Nakade K.,
RA   Nakamura Y., Nashimoto H., Nishio Y., Oyama S., Saito N.,
RA   Sampei G., Soki Y., Tagami H., Takemoto K., Wada C., Yamamoto Y.,
RA   Yano M., Horiuchi T.;
RT   "A 718-kb DNA sequence of the Escherichia coli K-12 genome
RT   corresponding to the 12.7-28.0 min region on the linkage map.";
RL   DNA Res. 3:137-155(1996).
RN   [3]
RP   SEQUENCE OF 1-92 FROM N.A.
RC   STRAIN K12 / W3110;
RX   MEDLINE=96276582; PubMed=7751275;
RA   Koh Y.-S., Roe J.-H.;
RT   "Isolation of a novel paraquat-inducible (pqi) gene-containing locus
RT   near ori2 in Escherichia coli.";
RL   J. Bacteriol. 177:2673-2678(1995).
CC   -!- INDUCTION: BY PARAQUAT.
CC   -!- SIMILARITY: BELONGS TO THE PQIB FAMILY.
CC   -!- ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is not modified
CC   and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -!- ----------------------------------------------------------------------
DR   EMBL; AE000197; AAC74032.1; -.
DR   EMBL; D90733; BAA15709.1; -.
DR   EMBL; X81561; CAA57257.1; -.
DR   HSSP; P17409; 1F2B.
DR   EcoGene; EG12966; pqiB.
KW   Transmembrane; Complete proteome.
FT   TRANSMEM     16     36       POTENTIAL.
SQ   SEQUENCE   546 AA;  60520 MW;  2C75CEB22ABBF7AB CRC64;


  Query Match             71.7%;  Score 33;  DB 1;  Length 546;
  Best Local Similarity   62.5%;  Pred. No. 38;
  Matches    5;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 YDDQESVK 8
           |||:|::
Db     271 YDDQKSIQ 278


RESULT  11
RAP1_YEAST
ID   RAP1_YEAST     STANDARD;      PRT;   827 AA.
AC   P11938;
DT   01-OCT-1989 (Rel. 12, Created)
DT   01-FEB-1995 (Rel. 31, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
DE   DNA-binding protein RAP1 (SBF E) (Repressor/activator site-binding
DE   protein) (TUF).
GN   RAP1 OR GRF1 OR TUF1 OR YNL216W OR N1410.
OS   Saccharomyces cerevisiae (Baker's yeast).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycetes; Saccharomycetales;
OC   Saccharomycetaceae; Saccharomyces.
OX   NCBI_TaxID=4932;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN BJ2168;
RX   MEDLINE=88052875; PubMed=3315241;
RA   Shore D., Nasmyth K.;
RT   "Purification and cloning of a DNA binding protein from yeast that
RT   binds to both silencer and activator elements.";
RL   Cell 51:721-732(1987).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN S288C / FY1679;
RX   MEDLINE=95282617; PubMed=7760814;
RA   Coster F., van Dyck L., Swaelen K., Purnelle B., Goffeau A.;
RT   "The sequence of a 24.1 kb DNA segment from the left arm of yeast
RT   chromosome XIV reveals MEF1, RAP1, a new putative member of the DNA
RT   replication complex and a new putative serine/threonine phosphatase
RT   gene.";
RL   Yeast 11:85-91(1995).
RN   [3]
RP   CHARACTERIZATION.
RX   MEDLINE=94254428; PubMed=8145855;
RA   Giraldo R., Rhodes D.;
RT   "The yeast telomere-binding protein RAP1 binds to and promotes the
RT   formation of DNA quadruplexes in telomeric DNA.";
RL   EMBO J. 13:2411-2420(1994).
RN   [4]
RP   X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS) OF 369-600.
RX   MEDLINE=96180725; PubMed=8620581;
RA   Konig P., Giraldo R., Chapman L., Rhodes D.;
RT   "The crystal structure of the DNA-binding domain of yeast RAP1 in
RT   complex with telomeric DNA.";
```

```
RL   Cell 85:12[ 18 (1996).
CC    !  FUNCTION: ESSENTIAL REGULATORY PROTEIN IN YEAST WHICH DNA BINDING
CC       SEQUENCE IS ONE OF THREE TYPES  F THE M.SPMA ELEMENTS IN MITOTIC
CC       SILENCING AND RECOMBINATION  RAP1 IS ALSO INVOLVED IN THE REGULATION
CC       OF TELOMERE STRUCTURE, WHERE ITS BINDING SITES ARE FOUND WITHIN
CC       THE TERMINAL POLY(1-3)A] SEQUENCES. THE OPPOSITE REGULATORY
CC       FUNCTIONS  RAP1 ASP %2 INTERACT  ITH ITS BINDING SITES DIF.
CC       INSTEAD, RESULT FROM INTERACTIONS WITH DIFFERENT FACTORS AT
CC       PROMOTERS AND SILENCERS. RAP1 ASSOCIATES WITH SIR3 AND SIR4
CC       PROTEINS TO FORM A DNA BINDING COMPLEX THAT INITIATES THE
CC       REPRESSION OF THE HM L AND TEL MEDIA  M/Y AI  ...
CC       BINDING OF RIF1 TO SILENCERS AND TELOMERES.
CC    !  SUBUNIT: RIF1, SIR3 AND SIR4 ASSOCIATE WITH RAP1 TERMINUS IN
CC       VIVO.
CC    !  SUBCELLULAR LOCATION: Nuclear.
CC    !  SIMILARITY: TO YEAST ARS BINDING FACTOR 1.
CC    !  SIMILARITY: CONTAINS 1 BRCT DOMAIN.
CC    !  SIMILARITY: CONTAINS 1 MYB LIKE DOMAIN.
CC    ---------- ---- -- ---- ---------- ---- ---------- ---- ---- ---- --
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------- ---- -- ---- ---------- ---- ---------- ---- ---- ---- --
DR   EMBL; M18068; AAA18404.1; -.
DR   EMBL; X78308; CAA55401.1; -.
DR   EMBL; Z71492; CAA95118.1; -.
DR   PIR; A29620; A29620.
DR   PDB; 1IGN; 27-JAN-97.
DR   TRANSFAC; T00715; -.
DR   SGD; S0005160; RAP1.
DR   InterPro; IPR001357; BRCT.
DR   InterPro; IPR001005; Myb_DNA_bind.
DR   Pfam; PF00533; BRCT; 1.
DR   Pfam; PF00249; myb_DNA_binding; 1.
DR   SMART; SM00292; BRCT; 1.
DR   SMART; SM00395; SANT; 1.
DR   PROSITE; PS50172; BRCT; 1.
DR   PROSITE; PS00037; MYB_1; FALSE_NEG.
DR   PROSITE; PS00334; MYB_2; FALSE_NEG.
DR   PROSITE; PS50090; MYB_3; 1.
KW   Transcription regulation; DNA binding; Telomere; Nuclear protein;
KW   Activator; Repressor; 3D-structure.
FT   DOMAIN      12    208       BRCT.
FT   DNA_BIND    355   411       MYB.
FT   DOMAIN      73    74        POLY-ASP.
FT   DOMAIN      83    84        POLY-THR.
FT   DOMAIN      317   320       POLY-ASN.
FT   DOMAIN      616   627       POLY-ALA.
FT   DOMAIN      63    695       ACTIVATION DOMAIN.
FT   CONFLICT    62    672       A -> R (IN REF. 1).
SQ   SEQUENCE   827 AA;  92413 MW;  67C4AFB04B8C0713 CRC64;


  Query Match             71.7%;  Score 33;  DB 1;  Length 8.7;
  Best Local Similarity   66.7%;  Pred. No. 60;
  Matches    6;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 YDHQESVKS 9
            :|| |:||
Db     789 HDDDESELKS 797


RESULT 12
IP2B_CHICK
ID   IP2B_CHICK     STANDARD;      PRT;  1627 AA.
AC   Q42131;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   DNA topoisomerase II, beta isozyme (EC 5.99.1.3).
GN   TOP2B.
OS   Gallus gallus (chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=9031;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Nitta A., Harata M., Mizuno S.;
RL   Submitted (??-???-????) to the EMBL/GenBank/DDBJ databases.
CC    !  FUNCTION: CONTROL OF TOPOLOGICAL STATES OF DNA BY TRANSIENT
CC       BREAKAGE AND SUBSEQUENT REJOINING OF DNA STRANDS. TOPOISOMERASE II
CC       MAKES DOUBLE-STRAND BREAKS (BY SIMILARITY).
CC    -!- CATALYTIC ACTIVITY: ATP-dependent breakage, passage and rejoining
CC       of double stranded DNA.
CC    !  SUBUNIT: HOMODIMER (BY SIMILARITY).
CC    !  SUBCELLULAR LOCATION: NUCLEAR; GENERALLY LOCATED IN THE
CC       NUCLEOPLASM.
CC    !  MISCELLANEOUS: EUKARYOTIC TOPOISOMERASE I AND II CAN RELAX BOTH
CC       NEGATIVE AND POSITIVE SUPERCOILS, WHEREAS IN EARLY 1? ENZYMES
CC       RELAX ONLY NEGATIVE SUPERCOILS.
CC    !  SIMILARITY: BELONGS TO THE TYPE II TOPOISOMERASE FAMILY.
CC    ---------- ---- -- ---- ---------- ---- ---------- ---- ---- ---- --
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------- ---- -- ---- ---------- ---- ---------- ---- ---- ---- --
DR   EMBL; AB007446; BAA22540.1; -.
DR   HSSP; P06786; 1BGW.
DR   InterPro; IPR003957; CBFA_NFYB_topis.
DR   InterPro; IPR001241; DNA_topoisoII.
DR   InterPro; IPR002205; DNA_topoisoIV.
DR   InterPro; IPR003594; HATPase_c.
DR   Pfam; PF00204; DNA_topoisoII; 1.
DR   Pfam; PF00521; DNA_topoisoIV; 1.
DR   Pfam; PF02518; HATPase_c; 1.
DR   PRINTS; PR00415; CCAATSUBUNIA.
DR   PRINTS; PR00418; TP12FAMILY.
DR   ProDom; PD000616; DNA_topoisoII; 1.
DR   SMART; SM00433; TOP2c; 1.
DR   SMART; SM00434; TOP4c; 1.
DR   PROSITE; PS00177; TOPOISOMERASE_II; 1.
KW   Isomerase; Topoisomerase; DNA-binding; ATP-binding; Nuclear protein.
FT   NP_BIND     187    192       ATP (POTENTIAL).
FT   ACT_SITE    831    831       DNA CLEAVAGE (BY SIMILARITY).
FT   DOMAIN        5      23      POLY-GLY.
FT   DOMAIN     1265   1268      POLY-LYS.
FT   DOMAIN     1388   1393      POLY-ASP.
FT   DOMAIN     1393   1396      POLY-ASN.
SQ   SEQUENCE   1627 AA;  183245 MW;  8B651D10A2CAD34B CRC64;


  Query Match             71.7%;  Score 33;  DB 1;  Length 1627;
  Best Local Similarity   66.7%;  Pred. No. 1.3e+02;
  Matches    6;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 YDHQESVKS 9
            :|| |:||
Db     548 YEDPESLKS 556


RESULT 13
VSP1_BPMV4
ID   VSP1_BPMV4     STANDARD;      PRT;   123 AA.
AC   Q04766;
DT   01-OCT-1993 (Rel. 27, Created)
DT   01-OCT-1993 (Rel. 27, Last sequence update)
DT   01-FEB-1994 (Rel. 28, Last annotation update)
```

```
DE   Structural protein (ORF1) (Fragment).
OS   Lactococcus delbrueckii bacteriophage mv4.
OC   Viruses.
OX   NCBI_TaxID=12892;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=93267750; PubMed=8497043;
RA   Vasala A., Dupont L., Baumann M., Ritzenthaler P., Alatossava T.;
RT   "Molecular comparison of the structural proteins encoding new
RT   clusters of two related Lactobacillus delbrueckii bacteriophages.";
RL   J. Virol. 67:3061-3068(1993).
CC   -!- SIMILARITY: TO BACTERIOPHAGE LL-H STRUCTURAL PROTEIN.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaborat
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a li cense agreement (See http://www.isb-sib.ch/ann
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; L32497; AAA56846.1; -.
DR   PIR; F45691; F45691.
KW   Structural protein.
FT   NON_TER       1       1
SQ   SEQUENCE   123 AA;  14600 MW;  17E5A0B3EEF7775E CRC64;


  Query Match             69.6%;  Score 32;  DB 1;  Length 123;
  Best Local Similarity   44.4%;  Pred. No. 11;
  Matches    4;  Conservative    5;  Mismatches    0;  Indels    4;  Gaps    0;

Qy       1 YDEQESVKS 9
           |:|:::::
Db      46 YDEQEAIRN 54


RESULT  14
HSLU_CAMJE
ID   HSLU_CAMJE     STANDARD;      PRT:   439 AA.
AC   Q9PHL0;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   ATP-dependent hsl protease ATP-binding subunit hslU.
GN   HSLU OR C1066_20.
OS   Campylobacter jejuni.
OC   Bacteria; Proteobacteria; epsilon subdivision; Campylobacter group;
OC   Campylobacter.
OX   NCBI_TaxID=197;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=NCTC 11168;
RX   MEDLINE=20558922; PubMed=10888204;
RA   Parkhill J., Wren B.W., Mungall K., Ketley J.M., Churcher C.,
RA   Basham D., Chillingworth T., Davies R.M., Feltwell T., Holroyd S.,
RA   Jagels K., Karlyshev A.V., Moule S., Pallen M.J., Penn C.W.,
RA   Quail M.A., Rajandream M.A., Rutherford K.M., van Vliet A.H.M.,
RA   Whitehead S., Barrell B.G.;
RT   "The genome sequence of the food-borne pathogen Campylobacter jejuni
RT   reveals hypervariable sequences.";
RL   Nature 403:665-668(2000).
CC   -!- FUNCTION: CHAPERONE SUBUNIT OF A PROTEASOME-LIKE DEGRADATION
CC       COMPLEX (BY SIMILARITY).
CC   -!- SUBUNIT: INTERACTS WITH HSLV (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC   -!- SIMILARITY: BELONGS TO THE CLPX CHAPERONE FAMILY. HSLU SUBFAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaborat
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announ
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AL139075; CAB75294.1; -.
DR   InterPro; IPR000593; AAA.
DR   InterPro; IPR000959; AAA_subfam.
DR   Pfam; PF00004; AAA; 1.
DR   SMART; SM00382; AAA; 1.
KW   Chaperone; ATP-binding; Complete proteome.
FT   NP_BIND      56      62       ATP (BY SIMILARITY).
SQ   SEQUENCE   439 AA;  49779 MW;  36FE3914DFBE48DA CRC64;


  Query Match             69.6%;  Score 32;  DB 1;  Length 439;
  Best Local Similarity   85.7%;  Pred. No. 47;
  Matches    6;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3 DQESVKS 9
           ||||:||
Db     230 DQESIKS 236


RESULT  15
AMY1_SACFI
ID   AMY1_SACFI     STANDARD;      PRT:   494 AA.
AC   P21567;
DT   01-MAY-1991 (Rel. 18, Created)
DT   01-MAY-1991 (Rel. 18, Last sequence update)
DT   01-APR-1994 (Rel. 25, Last annotation update)
DE   Alpha amylase precursor (EC 3.2.1.1).
GN   ALP1.
OS   Saccharomycopsis fibuligera (Yeast).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycetina; Saccharomycetes;
OC   Saccharomycetales; Saccharomycopsidaceae; Saccharomycopsis.
OX   NCBI_TaxID=4944;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=87279512; PubMed=3497657;
RA   Itoh T., Yamashita I., Fukui S.;
RT   "Nucleotide sequence of the alpha-amylase gene (ALP1) in the yeast
RT   Saccharomycopsis fibuligera.";
RL   FEBS Lett. 219:339-342(1987).
CC   -!- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-alpha-D-glucosidic
CC       linkages in oligosaccharides and polysaccharides.
CC   -!- COFACTOR: BINDS A CALCIUM ION REQUIRED FOR ITS ACTIVITY.
CC   -!- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, A.K.A.
CC       KNOWN AS THE ALPHA AMYLASE FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaborat
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announ
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X05791; CAA29283.1; -.
DR   PIR; S00064; ALBYAF.
DR   HSSP; P10529; 7TAA.
DR   InterPro; IPR000461; Alpha-amylase.
DR   Pfam; PF00128; alpha-amylase; 1.
DR   Pfam; PF02806; alpha-amylase_C; 1.
KW   Hydrolase; Glycosidase; Carbohydrate metabolism; Calcium.
KW   Glycoprotein; Signal.
FT   SIGNAL        1      26       POTENTIAL.
FT   CHAIN        27     494       ALPHA AMYLASE.
FT   ACT_SITE    233     233       BY SIMILARITY.
FT   ACT_SITE    237     237       BY SIMILARITY.
FT   ACT_SITE    324     324       BY SIMILARITY.
FT   CARBOHY:    224     224       N-LINKED (GLCNAC...) (POTENTIAL)
FT   DISULFID     57      65       BY SIMILARITY.
FT   DISULFID    177     191       BY SIMILARITY.
FT   DISULFID    267     410       BY SIMILARITY.
```

FI [illegible] BY SIMILARITY.
SQ SEQUENCE 494 AA; [illegible] MW; [illegible]

Query Match 69.8%; Score 32; DB 1; Length 494;
Best Local Similarity 66.7%; Pred. No. 54;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 YDQQESVKS 9
|||| |:|
Db 182 YDQQAQVQS 190

Search completed: August 26, 2002, 06:29:03
Job time: 929 sec

Geneseq version 4.5
Copyright (c) 1995 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 26, 2002, 06:28:05 ; Search time 130.19 Seconds
(without alignments)
11.959 Million cell updates/sec

Title: US-09-744-282-3
Perfect score: 46
Sequence: 1 YDDQESVKS 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 1.0%
Listing first 45 summaries

Database : SPTREMBL_19:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriap:*
17: sp_archaep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	46	100.0	383	4	Q96HI7	Q96hi7 homo sapien
2	43	93.5	332	6	Q18949	Q18949 bos taurus
3	43	93.5	383	6	Q29411	Q29411 sus scrofa
4	37	80.4	352	11	Q9WTV1	Q9wtv1 rattus norv
5	37	80.4	4095	2	Q02425	Q02425 caenorhabdi
6	36	78.3	274	4	Q99M65	Q99m65 homo sapien
7	36	78.3	1170	5	Q17490	Q17490 caenorhabdi
8	36	78.3	6994	5	Q17343	Q17343 caenorhabdi
9	35	76.1	235	10	Q97K67	Q97k67 streptococc
10	35	76.1	381	11	Q99J84	Q99j84 mus musculu
11	35	76.1	477	5	Q90620	Q90620 suberites d
12	35	76.1	819	5	Q95Q49	Q95q49 caenorhabdi
13	35	76.1	1586	13	Q9DFB7	Q9dfb7 gallus gall
14	34	73.9	441	2	Q60035	Q60035 thermotoga
15	34	73.9	674	2	Q9FAC8	Q9fac8 bacillus co
16	34	73.9	676	2	Q93AD3	Q93ad3 bacillus th
17	34	73.9	774	13	Q9DF43	Q9df43 gallus gall
18	34	73.9	886	4	Q9UMG8	Q9umg8 homo sapien
19	34	73.9	955	12	Q37163	Q37163 bluetongue
20	34	73.9	955	12	Q37164	Q37164 bluetongue
21	34	73.9	955	12	Q86782	Q86782 bluetongue
22	34	73.9	1724	5	Q9N583	Q9n583 caenorhabdi
23	34	73.9	1758	5	Q95Y84	Q95y84 caenorhabdi
24	33	71.7	374	16	Q9PJX0	Q9pjx0 chlamydia m
25	33	71.7	385	17	Q97CJ4	Q97cj4 thermoplasm
26	33	71.7	394	5	Q9N4W1	Q9n4w1 caenorhabdi
27	33	71.7	428	10	Q9LUV7	Q9luv7 arabidopsis
28	33	71.7	1058	10	O80486	O80486 arabidopsis
29	33	71.7	1075	3	Q01805	Q01805 pneumocysti
30	33	71.7	1088	3	Q07242	Q07242 pneumocysti
31	33	71.7	1199	3	Q01816	Q01816 pneumocysti
32	33	71.7	1248	10	O65589	O65589 arabidopsis
33	32	69.6	106	10	Q9FWP1	Q9fwp1 oryza sativ
34	32	69.6	176	16	Q9W2C4	Q9w2c4 thermotoga
35	32	69.6	230	5	O16265	O16265 caenorhabdi
36	32	69.6	324	11	Q9EQN8	Q9eqn8 mus musculu
37	32	69.6	369	2	Q56249	Q56249 thermophili
38	32	69.6	373	2	Q52100	Q52100 photobacter
39	32	69.6	413	2	O66145	O66145 citrobacter
40	32	69.6	429	5	Q95QA5	Q95qa5 caenorhabdi
41	32	69.6	480	3	Q9P874	Q9p874 candida alb
42	32	69.6	516	10	O48708	O48708 arabidopsis
43	32	69.6	603	5	Q95QA7	Q95qa7 caenorhabdi
44	32	69.6	652	5	Q07615	Q07615 plasmodium
45	32	69.6	663	5	Q95QA6	Q95qa6 caenorhabdi

ALIGNMENTS

RESULT 1
Q96HI7
ID Q96HI7 PRELIMINARY; PRT; 383 AA.
AC Q96HI7;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE SIMILAR TO CHITINASE 3-LIKE 1 (CARTILAGE GLYCOPROTEIN-39).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BRAIN, AND GLIOBLASTOMA WITH EGFR AMPLIFICATION;
RA Stransberg R.;
RL Submitted (MAY 2001) to the EMBL/GenBank/DDBJ databases.
DR EMBL; BC008568; AAH08568.1; -.
SQ SEQUENCE 383 AA; 42625 MW; 6C9EF134B8B7C2D1 CRC64;

Query Match 100.0%; Score 46; DB 4; Length 383;
Best Local Similarity 100.0%; Pred. No. 0.47;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YDDQESVKS 9
|||||||||
Db 328 YDDQESVKS 336

RESULT 2
Q18949
ID Q18949 PRELIMINARY; PRT; 332 AA.
AC Q18949;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE CHITINASE LIKE PROTEIN 1 (FRAGMENT).
GN CLP-1.

OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RA Recklies A.D., White C.;
RT "Expression of chitinase-like protein 1 (CLP-1) in bovine
RT chondrocytes.";
RL Submitted (JUN-1997) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AF011373; AAB64304.1; -.
DR InterPro; IPR001223; Glyco_hydro_18.
DR Pfam; PF00704; Glyco_hydro_18; 1.
FT NON_TER 1 1
FT NON_TER 332 332
SQ SEQUENCE 332 AA; 37434 MW; [illegible];

Query Match 93.5%; Score 43; DB 6; Length 332;
Best Local Similarity 88.9%; Pred. No. 1.3;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YDDQESVKS 9
|||:|||||
Db 304 YDDQESVKN 312

RESULT 3
Q29411
ID Q29411 PRELIMINARY; PRT; 393 AA.
AC Q29411;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE 38 KDA HEPARIN-BINDING GLYCOPROTEIN.
OS Sus scrota (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=SMOOTH MUSCLE;
RX MEDLINE=95286589; PubMed=7768902;
RA Shackelton L.M., Mann D.M., Millis A.J.;
RT "Identification of a 38-kDa heparin-binding glycoprotein (gp38k) in
RT differentiating vascular smooth muscle cells as a member of a group of
RT proteins associated with tissue remodeling.";
RL J. Biol. Chem. 270:13076-13083(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=SMOOTH MUSCLE;
RA Shackelton L.M., Mann D.M., Millis A.J.T.;
RL Submitted (JAN-1995) to the EMBL/GenBank/DDBJ databases.
DR EMBL; U19900; AAA86482.1; -.
DR EMBL; Z47803; CAA87764.1; -.
DR InterPro; IPR001223; Glyco_hydro_18.
DR Pfam; PF00704; Glyco_hydro_18; 1.
KW Heparin-binding.
SQ SEQUENCE 393 AA; 42443 MW; [illegible];

Query Match 93.5%; Score 43; DB 6; Length 393;
Best Local Similarity 88.9%; Pred. No. 1.5;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YDDQESVKS 9
|||||||||:
Db 328 YDDQESVKN 336

RESULT 4
Q9WTV1

ID Q9WTV1 PRELIMINARY; PRT; 352 AA.
AC Q9WTV1;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE GLYCOPROTEIN-39 (FRAGMENT).
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LEWIS;
RA Wendling U., Boots A.M.H., van Eden W.;
RT "Cloning of the rat homologue of human cartilage glycoprotein-39 a
RT potential autoantigen in arthritis.";
RL Submitted (AUG-1999) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AF062038; AAD22610.1; -.
DR HSSP; P07254; 1CTN.
DR InterPro; IPR000677; 2S_Globulin.
DR InterPro; IPR001223; Glyco_hydro_18.
DR Pfam; PF00704; Glyco_hydro_18; 1.
DR PRINTS; PR00551; 2SGLOBULIN.
FT NON_TER 1 1
FT NON_TER 352 352
SQ SEQUENCE 352 AA; 39991 MW; [illegible];

Query Match 80.4%; Score 37; DB 11; Length 352;
Best Local Similarity 77.8%; Pred. No. 22;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 YDDQESVKS 9
||| ||||:
Db 307 YDDPESVKN 315

RESULT 5
O02425
ID O02425 PRELIMINARY; PRT; 4063 AA.
AC O02425; O17952; Q22014; Q61612;
DT 01-JUL-1997 (TrEMBLrel. 04, Created)
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE SMA-1 PROTEIN (BETA CHAIN SPECTRIN HOMOLOG SMA1).
GN SMA-1 OR SMA1.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditina;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA Lennard N.;
RL Submitted (JUL-1996) to the EMBL/GenBank/DDBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99069613; PubMed=9851916;
RA none;
RT "Genome sequence of the nematode C.elegans: A platform for
RT investigating biology.";
RL Science 282:2012-2018(1998).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=98247708; PubMed=9570774;
RA McKeown C., Praitis V., Austin J.;
RT "sma-1 encodes a betaH-spectrin homolog required for Caenorhabditis
RT elegans morphogenesis.";
RL Development 125:2087-2098(1998).
DR EMBL; Z75956; CAB00130.2; -.
DR EMBL; Z81570; CAB00130.2; JOINED.
DR EMBL; Z81570; CAB04608.2; -.
DR EMBL; Z75956; CAB04608.2; JOINED.
DR EMBL; AF053496; AAC08577.1; -.

DR InterPro; IPR0 [illegible]; Actinin_actin_bind.
DR InterPro; IPR0 [illegible]; Calponin_hom.
DR InterPro; IPR0 [illegible]; PH.
DR InterPro; IPR006088; PFTA.
DR InterPro; IPR0 [illegible]; SH3.
DR InterPro; IPR0 [illegible]; Spectrin.
DR InterPro; IPR0 [illegible]; Spectrin_IH.
DR Pfam; PF00307; CH; 2.
DR Pfam; PF00169; PH; 1.
DR Pfam; PF00018; SH3; 1.
DR Pfam; PF00435; spectrin; 11.
DR PRINTS; PR00888; SPECTRINPH.
DR SMART; SM00033; CH; 2.
DR SMART; SM00233; PH; 1.
DR SMART; SM00326; SH3; 1.
DR SMART; SM00150; SPEC; 20.
DR PROSITE; PS00019; ACTININ_1; 1.
DR PROSITE; PS00020; ACTININ_2; 1.
DR PROSITE; PS50021; CH; 2.
DR PROSITE; PS50003; PH_DOMAIN; 1.
DR PROSITE; PS00904; PFTA; UNKNOWN_1.
DR PROSITE; PS50002; SH3; 1.
SQ SEQUENCE 4666 AA; 47006 MW; 1EC11D7DB0B83715 CRC64;

Query Match 80.4%; Score 47; DB 5; Length 4666;
Best Local Similarity 66.7%; Pred. No. 4.2e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 YDDQESVKS 9
|| :|:||
Db 4046 YDDSDSIKS 4054

RESULT 6
Q96M65
ID Q96M65 PRELIMINARY; PRT; 271 AA.
AC Q96M65;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE cDNA FLJ42742 FIS, CLONE TESTI2002369.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=TESTIS;
RA Ishibashi T., Kimchoji K., Yosida M., Watanabe S., Ishida S., Ono Y.,
RA Hotuta T., Hiraoka J., Murakawa K., Takiguchi S., Kusano J.,
RA Watanabe M., Fujimori K., Tanai H., Ishida M., Yamashita H., Chiba Y.,
RA Sugiyama T., Ito K., Otsuki T., Sato H., Wakamatsu A., Ishii S.,
RA Yamamoto J., Isogai Y., Kawai-Hio Y., Saito K., Nishikawa T.,
RA Kimura K., Matsuo K., Nakamura T., Sekito M., Aikoh H., Kondo K.,
RA Watatsuma M., Tarohchi Fujii A., Ochima A., Shiriyama A., Kazakami F.,
RA Suzuki Y., Sugano S., Nagahari K., Masuho Y., Nagai K., Isogai T.;
RT "NEDO human cDNA sequencing project.";
RL Submitted (OCT-2001) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AK057354; BAB71445.1; -.
SQ SEQUENCE 271 AA; 29401 MW; A1984F6AFD8F8F1E CRC64;

Query Match 78.3%; Score 36; DB 4; Length 271;
Best Local Similarity 75.0%; Pred. No. 26;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YDDQESVK 8
|:|:|||:
Db 171 YDDEESLK 178

RESULT 7
Q17490
ID Q17490 PRELIMINARY; PRT; 5170 AA.
AC Q17490;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE B0350.1 PROTEIN.
GN B0350.1.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91158718; PubMed=2006398;
RA Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,
RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,
RA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,
RA Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,
RA Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,
RA Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Shownkeen R.,
RA Smaldon N., Smith A., Sonnhammer E., Staden R., Sulston J.,
RA Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,
RA Watson A., Weinstock L., Wilkinson-Sproat J., Wohldman P.;
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT elegans.";
RL Nature 368:32-38(1994).
RN [2]
RP SEQUENCE FROM N.A.
RA Gattung S.;
RT "The sequence of C. elegans cosmid B0350.";
RL Submitted (MAR-1996) to the EMBL/GenBank/DDBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Waterston R.;
RL Submitted (FEB-1996) to the EMBL/GenBank/DDBJ databases.
DR EMBL; U50071; AAA93447.1; -.
SQ SEQUENCE 5170 AA; 575848 MW; 27C2DFBF4AE03A7B CRC64;

Query Match 78.3%; Score 36; DB 5; Length 5170;
Best Local Similarity 87.5%; Pred. No. 6.6e+02;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 DDQESVKS 9
|||| :||
Db 2119 DDQESLKS 2126

RESULT 8
Q17343
ID Q17343 PRELIMINARY; PRT; 6994 AA.
AC Q17343;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE UNC-44 ANKYRINS.
GN UNC-44.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN N2;
RX MEDLINE=95263663; PubMed=7744957;
RA Otsuka A.J., Franco R., Yang B., Shim K., Tang L.Z., Zhang Y.,
RA Boontrakulpoontawee P., Jeyaprakash A., Hedgecock E., Wheaton V.I.,
RA Sobery A.;
RT "An ankyrin related gene (unc-44) is necessary for proper axonal
RT guidance in Caenorhabditis elegans.";

```
RL   J. Cell Biol. 129:1081-1092(1995).
RN   [2]
RP   REVISIONS, AND SEQUENCE OF 6126-6994 FROM N.A.
RC   STRAIN=N2;
RA   Otsuka A.J.;
RL   Submitted (SEP-1996) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U39847; AAB41827.1; -.
DR   EMBL; U21733; AAB38384.1; -.
DR   HSSP; P42774; 1IHB.
DR   InterPro; IPR002110; ANK.
DR   InterPro; IPR000488; Death.
DR   InterPro; IPR002383; CLA_blood.
DR   InterPro; IPR001460; Glyco_hydro_1.
DR   InterPro; IPR000906; ZU5.
DR   Pfam; PF00023; ank; 24.
DR   Pfam; PF00531; death; 1.
DR   Pfam; PF00791; ZU5; 1.
DR   PRINTS; PR01415; ANKYRIN.
DR   PRINTS; PR00031; GLABLOOD.
DR   SMART; SM00248; ANK; 21.
DR   SMART; SM00005; DEATH; 1.
DR   SMART; SM00218; ZU5; 1.
DR   PROSITE; PS50088; ANK_REPEAT; 22.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
DR   PROSITE; PS50017; DEATH_DOMAIN; 1.
DR   PROSITE; PS00572; GLYCOSYL_HYDROL_F1_1; UNKNOWN_2.
KW   ANK repeat; Repeat.
SQ   SEQUENCE   6994 AA;  775364 MW;  90CB449925D992BD CRC64;


  Query Match             78.3%;  Score 46;  DB 5;  Length 6994;
  Best Local Similarity   87.5%;  Pred. No. 9.2e+02;
  Matches    7;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2 DDQESVKS 9
            |||||:||
Db    3943 DDQESLES 3950


RESULT   9
Q97R37
ID   Q97R37       PRELIMINARY;      PRT;   235 AA.
AC   Q97R37;
DT   01-OCT-2001 (TrEMBLrel. 18, Created)
DT   01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   CYTOCHROME C-TYPE BIOGENESIS PROTEIN CCDA.
GN   SP0999.
OS   Streptococcus pneumoniae.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC   Streptococcus.
OX   NCBI_TaxID=1313;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=TIGR4;
RX   MEDLINE=21357209; PubMed=11463916;
RA   Tettelin H., Nelson K.E., Paulsen I.T., Eisen J.A., Read T.D.,
RA   Peterson S., Heidelberg J., DeBoy R.T., Haft D.H., Dodson R.J.,
RA   Durkin A.S., Gwinn M., Kolonay J.F., Nelson W.C., Peterson J.D.,
RA   Umayam L.A., White O., Salzberg S.L., Lewis M.R., Radune D.,
RA   Holtzapple E., Khouri H., Wolf A.M., Utterback T.R., Hansen C.L.,
RA   McDonald L.A., Feldblyum T.V., Angiuoli S., Dickinson T., Hickey E.K.,
RA   Holt I.E., Loftus B.J., Yang F., Smith H.O., Venter J.C.,
RA   Dougherty B.A., Morrison D.A., Hollingshead S.K., Fraser C.M.;
RT   "Complete genome sequence of a virulent isolate of Streptococcus
RT   pneumoniae.";
RL   Science 293:498-506(2001).
DR   EMBL; AF334183; AAK75116.1; -.
DR   TIGR; SP0999; .
DR   InterPro; IPR001128; Cyt_P450.
DR   InterPro; IPR000833; DsbD_DipZ.
DR   Pfam; PF02683; DsbD; 1.
DR   PROSITE; PS00086; CYTOCHROME_P450; UNKNOWN_1.
KW   Complete proteome.
SQ   SEQUENCE   235 AA;  25898 MW;  E6CB5C9871C33AA2 CRC64;


  Query Match             76.1%;  Score 45;  DB 16;  Length 235;
  Best Local Similarity   87.5%;  Pred. No. 46;
  Matches    7;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       2 DDQESVKS 9
            |||||:||
Db      36 DDQESAKS 43


RESULT  10
Q99J84
ID   Q99J84       PRELIMINARY;      PRT;   381 AA.
AC   Q99J84;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-OCT-2001 (TrEMBLrel. 18, Last annotation update)
DE   SIMILAR TO CHITINASE 3-LIKE 1 (CARTILAGE GLYCOPROTEIN-39).
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Strausberg R.;
RL   Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Strausberg R.;
RL   Submitted (FEB-2001) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RA   Strausberg R.;
RL   Submitted (MAR-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC005611; AAH05611.1; -.
DR   EMBL; BC003786; AAH03786.1; -.
DR   EMBL; BC004734; AAH04734.1; -.
DR   InterPro; IPR001223; Glyco_hydro_18.
DR   Pfam; PF00704; Glyco_hydro_18; 1.
SQ   SEQUENCE   381 AA;  42979 MW;  EF508C9AE9D445C0 CRC64;


  Query Match             76.1%;  Score 45;  DB 11;  Length 381;
  Best Local Similarity   66.7%;  Pred. No. 60;
  Matches    6;  Conservative    3;  Mismatches    0;  Indels    1;  Gaps    0;

Qy       1 YDDQESVKS 9
            |:|:|||:|
Db     329 YEDKESVKN 337


RESULT  11
Q966Z0
ID   Q966Z0       PRELIMINARY;      PRT;   477 AA.
AC   Q966Z0;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   CORTACIIN.
OS   Suberites domuncula.
OC   Eukaryota; Metazoa; Porifera; Demospongiae; Tetractinomorpha;
OC   Hadromerida; Suberitidae; Suberites.
OX   NCBI_TaxID=55567;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Mueller W.E.G., Schatton W., Wimmer W., Bohm M., Bater R., Rau F.,
RT   "Initiation of an aquaculture of sponges for their sustainable
RT   production of bioactive metabolites in open systems.";
RL   Mol. Mar. Biol. Biotechnol. 1:569-579(1999).
DR   EMBL; Y18027; CAC48778.1; -.
```

```
SQ   SEQUENCE   477 AA;   2549 MW;   ID6D6042F50AD2C7 CRC64;


  Query Match          76.1%;  Score 35;  DB 5;  Length 477;
  Best Local Similarity   66.7%;  Pred. No. 77;
  Matches    6;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 YDDQESVKS 9
           |:: :|||
Db     288 YDDMQAVES 296


RESULT  12
Q95Q49
ID   Q95Q49      PRELIMINARY;      PRT;   819 AA.
AC   Q95Q49;
DT   01 DEC 2001 (TrEMBLrel. 19, Created)
DT   01 DEC 2001 (TrEMBLrel. 19, Last sequence update)
DT   01 DEC 2001 (TrEMBLrel. 19, Last annotation update)
DE   HYPOTHETICAL 92.5 KDA PROTEIN.
GN   F56B6.2.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN BRISTOL N2;
RX   MEDLINE=98006861; PubMed=9851916;
RA   None;
RT   "Genome sequence of the nematode C. elegans: a platform for
RT   investigating biology. The C. elegans Sequencing Consortium.";
RL   Science 282:2012-2018(1998).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN BRISTOL N2;
RA   Stelyes L., Stelyes L.;
RT   "The sequence of C. elegans cosmid F56B6.";
RL   Submitted (JUL-1999) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN BRISTOL N2;
RA   Waterston R.;
RT   "Direct Submission.";
RL   Submitted (OCT-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U64599; AAL11462.1; -.
KW   Hypothetical protein.
SQ   SEQUENCE   819 AA;  92589 MW;  A84B6B5F805CC6B5 CRC64;


  Query Match          76.1%;  Score 35;  DB 5;  Length 819;
  Best Local Similarity   66.7%;  Pred. No. 1.4e+02;
  Matches    6;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 YDDQESVKS 9
           |||| |: |
Db       9 YDDQSSISS 17


RESULT  13
Q9DEB7
ID   Q9DEB7      PRELIMINARY;      PRT;  1586 AA.
AC   Q9DEB7;
DT   01 MAR 2001 (TrEMBLrel. 16, Created)
DT   01 MAR 2001 (TrEMBLrel. 16, Last sequence update)
DT   01 DEC 2001 (TrEMBLrel. 19, Last annotation update)
DE   TOPOISOMERASE II ALPHA 2.
GN   TOP2A.
OS   Gallus gallus (chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=9031;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Petruti-Mot A.S., Earnshaw W.C.;
RT   "Two differentially spliced forms of topoisomerase II alpha and beta
RT   mRNAs are conserved between birds and humans.";
RL   Submitted (JUL-2000) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: CONTROL OF TOPOLOGICAL STATES OF DNA BY TRANSIENT
CC       BREAKAGE AND SUBSEQUENT RESEALING OF DNA STRANDS. TOPOISOMERASE II
CC       MAKES DOUBLE-STRAND BREAKS (BY SIMILARITY).
CC   -!- CATALYTIC ACTIVITY: ATP DEPENDENT BREAKAGE, PASSAGE AND RESEALING
CC       OF DOUBLE-STRANDED DNA.
CC   -!- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC   -!- MISCELLANEOUS: EUKARYOTIC TOPOISOMERASE I AND II CAN RELAX BOTH
CC       NEGATIVE AND POSITIVE SUPERCOILS, WHEREAS PROKARYOTIC ENZYMES
CC       RELAX ONLY NEGATIVE SUPERCOILS (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE TYPE II TOPOISOMERASE FAMILY.
DR   EMBL; AF285155; AAG13401.1; -.
DR   HSSP; P06786; 1BGW.
DR   InterPro; IPR003957; CBFA_NFYB_topis.
DR   InterPro; IPR001241; DNA_topoisoII.
DR   InterPro; IPR002205; DNA_topoisoIV.
DR   InterPro; IPR003594; HATPase_c.
DR   Pfam; PF00204; DNA_topoisoII; 2.
DR   Pfam; PF00521; DNA_topoisoIV; 1.
DR   Pfam; PF02518; HATPase_c; 1.
DR   PRINTS; PR00615; CCAATSUBUNTA.
DR   PRINTS; PR00416; TOP2FAMILY.
DR   ProDom; PD000616; DNA_topoisoII; 1.
DR   SMART; SM00387; HATPase_c; 1.
DR   SMART; SM00433; TOP2c; 1.
DR   SMART; SM00434; TOP4c; 1.
DR   PROSITE; PS00177; TOPOISOMERASE_II; 1.
KW   ATP-binding; DNA-binding; Isomerase; Topoisomerase.
SQ   SEQUENCE   1586 AA;  178839 MW;  422CF3AAC60EA8B6 CRC64;


  Query Match          76.1%;  Score 35;  DB 13;  Length 1586;
  Best Local Similarity   66.7%;  Pred. No. 2.9e+02;
  Matches    6;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 YDDQESVKS 9
           |:|:|||:|
Db     558 YEDRESLKS 566


RESULT  14
Q60035
ID   Q60035      PRELIMINARY;      PRT;   441 AA.
AC   Q60035;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   4-ALPHA-GLUCANOTRANSFERASE (EC 2.4.1.25).
GN   MGT.
OS   Thermotoga maritima.
OC   Bacteria; Thermotogales; Thermotoga.
OX   NCBI_TaxID=2336;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN MSB8 (DSM3109);
RX   MEDLINE=92341687; PubMed=1628664;
RA   Liebl W., Feil R., Gabelsberger J., Kellermann J., Schleifer K.;
RT   "Purification and characterization of a novel thermostable 4 alpha
RT   glucanotransferase of Thermotoga maritima cloned in Escherichia
RT   coli.";
RL   Eur. J. Biochem. 207:81-88(1992).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN MSB8 (DSM3109);
RA   Heinrich P., Huber W., Liebl W.;
RT   "Expression in Escherichia coli and structure of the gene encoding 4
RT   alpha-glucanotransferase from Thermotoga maritima. Classification of
RT   maltodextrin glycosyltransferases into two distantly related ...";
```

```
RL   Syst. Appl. Microbiol. 17:297-305(1994).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN MSBB (DSM4109);
RA   Liebl W.;
RL   Submitted (AUG 1995) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; Z50813; CAA90693.1; -.
DR   InterPro; IPR000461; Alpha_amylase.
DR   Pfam; PF00128; alpha_amylase; 1.
KW   Transferase; Glycosyltransferase.
SQ   SEQUENCE   447 AA;  51857 MW;  8D221B5234BAFBAC CRC64;


  Query Match             73.9%;  Score 34;  DB 2;  Length 447;
  Best Local Similarity   75.0%;  Pred. No. 1.1e+02;
  Matches    6;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 YDDQESVK 8
           |||| |:|
Db     409 YDDQHSLK 416


RESULT  15
Q9FAC8
ID   Q9FAC8       PRELIMINARY;      PRT;   674 AA.
AC   Q9FAC8;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   CHITINASE B.
GN   CHIB.
OS   Bacillus cereus.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group;
OC   Bacillus/Staphylococcus group; Bacillus.
OX   NCBI_TaxID=1396;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CH;
RA   Maburchi N., Araki Y.;
RT   "Chitinase.";
RL   Submitted (APR-2000) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: CONTAINS 1 FIBRONECTIN TYPE III-LIKE DOMAIN.
DR   EMBL; AB041932; BAB16891.1; -.
DR   HSSP; P07254; 1CTN.
DR   InterPro; IPR001919; CBD_2.
DR   InterPro; IPR003961; FN_III.
DR   InterPro; IPR001223; Glyco_hydro_18.
DR   InterPro; IPR001412; tRNA synt_1.
DR   Pfam; PF00553; CBD_2; 1.
DR   Pfam; PF00041; fn3; 1.
DR   Pfam; PF00704; Glyco_hydro_18; 1.
DR   SMART; SM00060; FN3; 1.
DR   PROSITE; PS00178; AA_TRNA_LIGASE_1; UNKNOWN_1.
KW   Hydrolase.
SQ   SEQUENCE   674 AA;  74262 MW;  66BE9FEB0D680561 CRC64;


  Query Match             73.9%;  Score 34;  DB 2;  Length 674;
  Best Local Similarity   75.0%;  Pred. No. 1.8e+02;
  Matches    6;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 YDDQESVK 8
           ||| |:|
Db     421 YDDNESMK 428



Search completed: August 26, 2002, 06:28:07
Job time: 1033 sec
```

Geneseq version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 26, 2002, 06:13:24 ; Search time 139.28 Seconds
(without alignments)
7.177 Million cell updates/sec

Title: US-09-744-282-3
Perfect score: 46
Sequence: 1 SDDQESVKES 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 241574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 241574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_082002:*
1: /SIP1/ptodata/1/pubpa/AA/AA_main.seq:*
2: /SIP1/ptodata/1/pubpa/AA/AA_new.seq:*
3: /SIP1/ptodata/1/pubpa/AA/AA_upd.seq:*
4: /SIP1/ptodata/1/pubpa/AA/AA_upd2.seq:*
5: /SIP1/ptodata/1/pubpa/AA/AA024.DAT:*
6: /SIP1/ptodata/1/pubpa/AA/AA1985.DAT:*
7: /SIP1/ptodata/1/pubpa/AA/AA_upd3.seq:*
8: /SIP1/ptodata/1/pubpa/AA/AA1995.DAT:*
9: /SIP1/ptodata/1/pubpa/AA/AA1996.DAT:*
10: /SIP1/ptodata/1/pubpa/AA/AA1997.DAT:*
11: /SIP1/ptodata/1/pubpa/AA/AA1998.DAT:*
12: /SIP1/ptodata/1/pubpa/AA/AA1999.DAT:*
13: /SIP1/ptodata/1/pubpa/AA/AA2000.DAT:*
14: /SIP1/ptodata/1/pubpa/AA/AA2001.DAT:*
15: /SIP1/ptodata/1/pubpa/AA/AA2002.DAT:*
16: /SIP1/ptodata/1/pubpa/AA/AA_upd4.seq:*
17: /SIP1/ptodata/1/pubpa/AA/AA_upd5.seq:*
18: /SIP1/ptodata/1/pubpa/AA/AA_upd6.seq:*
19: /SIP1/ptodata/1/pubpa/AA/AA_upd7.seq:*
20: /SIP1/ptodata/1/pubpa/AA/AA_upd8.seq:*
21: /SIP1/ptodata/1/pubpa/AA/AA_upd9.seq:*
22: /SIP1/ptodata/1/pubpa/AA/AA_upd10.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	46	100.0	9	12	AAB95073 Human cartilage gl
2	46	100.0	9	21	AAY69228 Peptide fragment o
3	46	100.0	13	17	AAB95077 Peptide coding for
4	46	100.0	14	21	AAY69252 Peptide fragment o
5	46	100.0	16	18	AAW45150 Human cartilage gl
6	46	100.0	16	18	AAW45151 Human cartilage gl
7	46	100.0	105	22	AAU02074 Synthetic human ti
8	46	100.0	220	22	AAU02078 Synthetic human mu
9	46	100.0	281	22	AAU02079 Synthetic multi-ta
10	46	100.0	821	4	AAP81842 Polypeptide involv
11	46	100.0	950	22	AAU02076 Synthetic human mu
12	46	100.0	459	22	AAU02077 Synthetic multi-ta
13	44	95.5	383	18	AAW26751 Bovine whey protei
14	44	95.5	383	21	AAB04442 Q-88k protein seq
15	42	91.3	14	22	ABB55896 Vascular dementia
16	42	91.3	14	22	ABB55940 Vascular dementia
17	42	91.3	14	22	ABB56115 Vascular dementia
18	42	91.3	14	22	ABB52309 Human APLP41 tryp
19	42	91.3	14	22	AAU28547 14I tryptic digest
20	42	91.3	14	22	AAU28750 16I tryptic digest
21	42	91.3	14	22	AAU26184 Depression Associa
22	42	91.3	14	22	AAU26398 Depression Associa
23	34	73.9	216	22	AAB63816 Human prostate can
24	34	73.9	1139	22	AAG22397 Novel human diagno
25	33	71.7	85	21	AAY96193 Protein encoded by
26	33	71.7	106	22	AAG82299 S. epidermidis ope
27	33	71.7	123	20	AAY47811 Chlamydia trachoma
28	33	71.7	1228	21	AAG30727 Arabidopsis thalia
29	33	71.7	1230	21	AAG30726 Arabidopsis thalia
30	33	71.2	1240	21	AAG30725 Arabidopsis thalia
31	32	69.6	244	22	AAG81736 S. epidermidis ope
32	32	69.6	452	15	AAK63186 Variant alpha amyl
33	32	69.6	468	13	AAR24136 Alpha-amylase vari
34	32	69.6	468	15	AAK63187 Variant alpha amyl
35	32	69.6	468	15	AAK63184 Variant alpha amyl
36	32	69.6	468	15	AAK63185 Variant alpha amyl
37	32	69.6	494	8	AAP70571 Alpha amylase gene
38	32	69.6	867	22	ABB61590 Drosophila melanog
39	31	67.4	188	21	AAB42977 Human CREX-CREZ741
40	31	67.4	190	22	ABG09331 Novel human diagno
41	31	67.4	243	22	ABG09332 Novel human diagno
42	31	67.4	329	21	AAG10436 Arabidopsis thalia
43	31	67.4	342	21	AAG10435 Arabidopsis thalia
44	31	67.4	358	21	AAG10434 Arabidopsis thalia
45	31	67.4	371	19	AAB76622 Human Bcl2-like pr

ALIGNMENTS

RESULT 1
AAB95073
ID AAB95073 standard; peptide; 9 AA.
XX
AC AAB95073;
XX
DT 17-JAN-1996 (first entry)
XX
DE Human cartilage glycoprotein 39 autoantigen derived peptide (111).
XX
KW human cartilage glycoprotein 39; hc gp-39; autoantigen;
KW cartilage repair; tolerance mediated; destruction;
KW rheumatoid arthritis, disease modify, diagnosis.
XX
OS Homo sapiens.
XX
PN WO9613517-A1.
XX
PD 09-MAY-1996.
XX
PF 25-OCT-1995; 95WO-EP04201.
XX
PR 07-APR-1995; 95EP-0200886.
PR 27-OCT-1994; 94EP-0203128.
XX
PA (ALKU) AKZO NOBEL NV.
XX
PI Boots AMB, Verheijden GFM;
XX
DR WPI; 1996-249447/24.
XX
PT Peptide(s) derived from human cartilage gp 39 autoantigen used
PT to induce specific T cell tolerance, esp. for treatment of
PT rheumatoid arthritis, also useful in producing arthritic animal

PT models
XX
PS Claim 1; Page 28; 37pp; English.
XX
CC The present peptide, derived from the human cartilage glycoprotein
CC 39 (hC gp 39) autoantigen, can be used to induce specific T cell
CC tolerance to hC gp 39 in a patient with T cell mediated cartilage
CC destruction, esp. rheumatoid arthritis (RA). It may also be used to
CC induce RA in mice, etc. (human disease models), and to diagnose the
CC presence of activated T cells autoreactive with hC gp 39, i.e. to
CC identify patients suitable for tolerising treatment.
CC Therapeutically the peptide is given in a dose of 0.1-100 microg/kg
CC by injection, or 10-50 microg/kg to produce a disease model. The
CC peptide has a specific tolerising effect on autoreactive T cells,
CC but leaves other components of the immune system intact, and
CC therefore should have few side effects. In an assay to determine
CC the peptide's binding to HLA DR mols., specifically DR4Rw4 (type 4
CC to be associated with an increased risk of RA development), the
CC peptide gave a DeltaG(50) (micromolar) value of 0.23 compared
CC to 0.55 for the control peptide HA(307-319).
XX
SQ Sequence 9 AA;

Query Match 100.0%; Score 46; DB 17; Length 9;
Best Local Similarity 100.0%; Pred. No. 6.4e-05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YDDQESVKS 9
||||||||| |
Db 1 yddqesvks 9

RESULT 2
AAY69228
ID AAY69228 standard; peptide; 9 AA.
XX
AC AAY69228;
XX
DT 30-MAY-2003 (first entry)
XX
DE Peptide fragment of human cartilage gp 39 (HC gp 39) protein.
XX
KW Human cartilage gp 39 protein; HC gp 39 protein; T-cell activator;
KW inflammatory disease; lymphocyte; antigen; bystander suppression;
KW antigen specific T cell tolerance; autoimmune disorder;
KW inflammatory disease; rheumatoid arthritis;
KW autoimmune inflammatory disease; autoantigen.
XX
OS Homo sapiens.
XX
PN WO200004917-A1.
XX
PD 03-FEB-2000.
XX
PF 19-JUL-1999; 99WO-EP05381.
XX
PR 23-JUL-1998; 98EP-0202471.
XX
PA (AKZO) AKZO NOBEL NV.
XX
PI Miltenburg AMM, Boots AMH;
XX
DR WPI; 2000-182587/16.
XX
PT Use of cartilage protein or peptides for treating inflammatory diseases
PT especially rheumatoid arthritis -
XX
PS Claim 2; Page 17; 29pp; English.
XX
CC AAY69226-33 represent fragments of the human cartilage gp-39 (HC gp-39)
CC protein. The peptides are T-cell activators. The peptides are used for
CC preventing inflammatory diseases, by modulating lymphocytes to antige-
CC to antigens other than HC gp 39, present in the same tissue as
CC HC gp 39. By the induction of antigen specific T cells, certain
CC autoimmune disorders can be treated by bystander suppression. The
CC HC gp 39 peptides are useful for preventing inflammatory diseases,
CC especially rheumatoid arthritis. The fragments may also be used for
CC treating autoimmune inflammatory diseases. The HC gp 39 peptides are
CC especially beneficial in autoimmune conditions in which the
CC autoantigens associated with the condition are unknown.
XX
SQ Sequence 9 AA;

Query Match 100.0%; Score 46; DB 21; Length 9;
Best Local Similarity 100.0%; Pred. No. 6.4e-05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YDDQESVKS 9
||||||||| |
Db 1 yddqesvks 9

RESULT 3
AAK95077
ID AAK95077 standard; peptide; 13 AA.
XX
AC AAK95077;
XX
DT 17-DEC-1996 (first entry)
XX
DE Peptide contg. hC gp 39 autoantigen derived peptide (III).
XX
KW Human; cartilage; glycoprotein 39; hC gp 39; autoantigen;
KW induction; T cell; tolerance; mediated; destruction;
KW rheumatoid arthritis; disease model; diagnosis.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Peptide 1..13
FT /note "HC gp 39 peptide (III)"
XX
PN WO9613517-A1.
XX
ID 09-MAY-1996.
XX
PF 25-OCT-1995; 95WO-EP04201.
XX
PR 07-APR-1995; 95EP-0200868.
PR 27-OCT-1994; 94EP-0202512.
XX
PA (AKZO) AKZO NOBEL NV.
XX
PI Boots AMH, Verheijden GFM;
XX
DR WPI; 1996-23944?/24.
XX
PT Peptide(s) derived from human cartilage gp 39 autoantigen, for
PT to induce specific T cell tolerance, esp. for treatment of
PT rheumatoid arthritis, also useful in producing arthritic animal
PT models
XX
PS Claim 2; Page 28; 37pp; English.
XX
CC The present peptide, contg. a peptide derived from the human
CC cartilage glycoprotein 39 (hC gp 39) autoantigen, can be used
CC induce specific T cell tolerance to hC gp 39 in a patient with
CC T-cell mediated cartilage destruction, esp. rheumatoid arthritis
CC (RA). It may also be used to induce RA in mice, etc. (human disease
CC models), and to diagnose the presence of activated T cells
CC autoreactive with hC gp 39, i.e. to identify patients suitable for
CC tolerising treatment.
CC Therapeutically the peptide is given in a dose of 0.1-100 microg/kg
CC by injection, or 10-50 microg/kg to produce a disease model. The

```
XX   peptide has a specific tolerising effect on autoreactive T cells,
XX   but leaves other long term's of the immune system intact, and
XX   therefore should have few side effects.
XX
SQ   Sequence   13 AA


  Query Match             100.0%;  Score 46;  DB 17;  Length 13;
  Best Local Similarity   100.0%;  Pred. No. 0.012;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 YDDQESVKS 9
           |||||||||
Db       5 yddqesvks 13


RESULT   4
AAY66242
ID   AAY66242 standard; peptide; 15 AA.
XX
AC   AAY66242;
XX
DT   06 MAY 2000  (first entry)
XX
DE   Peptide fragment of human cartilage gp-39 (HC gp 39) protein.
XX
KW   Human cartilage gp 39 protein; HC gp 39 protein; T-cell activator;
KW   inflammatory disease; lymphocyte; antigen; bystander suppression;
KW   antigen specific T cell; tolerance; autoimmune disorder;
KW   inflammatory disease; rheumatoid arthritis;
KW   autoimmune inflammatory disease; autoantigen.
XX
OS   Homo sapiens.
XX
PN   WO200004917-A2.
XX
PD   03-FEB-2000.
XX
PF   19-JUL-1999;   99WO-EP05351.
XX
PR   23-JUL-1998;   98EP-0202471.
XX
PA   (ALKO ) AKZO NOBEL NV
XX
PI   Mitchelmore AMM;  Boots AMH;
XX
DR   WPI; 2000-162172/14.
XX
PT   Use of cartilage protein or peptides for treating inflammatory diseases
PT   especially rheumatoid arthritis
XX
PS   Claim 2; Page 17; 29pp; English.
XX
CC   AAY66236-55 represent fragments of the human cartilage gp-39 (HC gp-39)
CC   protein. The peptides are T-cell activators. The peptides are used for
CC   preventing inflammatory diseases, by modulating lymphocytes reactive
CC   to antigens other than HC gp 39, present in the same tissue of
CC   HC gp 39. By the induction of antigen-specific T cell tolerance,
CC   autoimmune disorders can be treated by bystander suppression. The
CC   HC gp 39 peptides are useful for preventing inflammatory diseases
CC   especially rheumatoid arthritis. The fragments may also be used for
CC   treating autoimmune inflammatory diseases. The HC gp 39 peptides are
CC   especially beneficial in autoimmune conditions in which the
CC   autoantigens associated with the condition are unknown.
XX
SQ   Sequence   15 AA;


  Query Match             100.0%;  Score 46;  DB 21;  Length 15;
  Best Local Similarity   100.0%;  Pred. No. 0.012;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 YDDQESVKS 9
           |||||||||
Db       3 yddqesvks 11


RESULT   5
AAW45150
ID   AAW45150 standard; peptide; 16 AA.
XX
AC   AAW45150;
XX
DT   28-APR-1998  (first entry)
XX
DE   Human cartilage glycoprotein 39 derived peptide #51.
XX
KW   Articular cartilage; immunosuppressive therapy; antigen; autoantigen;
KW   immunological tolerance; T cell; human cartilage glycoprotein 39;
KW   HC gp-39; rheumatoid arthritis; epitope.
XX
OS   Synthetic.
OS   Homo sapiens.
XX
PN   WO9740068-A1.
XX
PD   30-OCT-1997.
XX
PF   22-APR-1997;   97WO-EP02051.
XX
PR   24-APR-1996;   96EP-0201106.
XX
PA   (ALKO ) AKZO NOBEL NV.
XX
PI   Boots AMH,  Verheijden GPM;
XX
DR   WPI; 1997-543775/49.
XX
PT   Peptide suitable for use in antigen specific immunosuppressive
PT   therapy - resembles or mimics epitope present on HC gp-39, so
PT   inducing systemic immunological tolerance to rheumatoid arthritis
PT   auto-antigen
XX
PS   Disclosure; Page 17; 82pp; English.
XX
CC   The present sequence represents a peptide which resembles or mimics an
CC   epitope present on human cartilage glycoprotein 39 (HC gp-39), an
CC   autoantigen in rheumatoid arthritis. The invention relates to peptides
CC   consisting of 16-55 amino acid residues comprising at least one of the
CC   following 15 sequences: AVVIASWG; FLGTNIPL; TISTAWIS; ASIMSVSS;
CC   FIKSVPPFL; FDGLDLAWL; LYPGRRDKQ; YDIAKISQH; LDFISIMTY; FISIMIYDF;
CC   PRGQEDASP; YAVGYMLKL; MLKIGAFAS; LAYYETCDF; IGGAIVHKI; YLKDKQCAI;
CC   LAGAMVWAL; VWALDLDDF; or LDLDDFQGS. They can be used medically in
CC   antigen specific immunosuppressive therapy, particularly the treatment
CC   of T-cell mediated destruction of articular cartilage in autoimmune
CC   diseases (e.g. rheumatoid arthritis). They can also be used to detect
CC   activated autoreactive T cells in an individual. The peptides have a
CC   specific effect on the autoreactive T cells, thus leaving the other
CC   components of the immune system intact, unlike the non specific
CC   suppressive effect of immunosuppressive drugs, and do not cause toxic
CC   side effects. The peptides are predominantly recognised by autoreactive
CC   T cells from rheumatoid arthritis patients, but rarely by those from
CC   healthy donors.
XX
SQ   Sequence   16 AA;


  Query Match             100.0%;  Score 46;  DB 18;  Length 16;
  Best Local Similarity   100.0%;  Pred. No. 0.015;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 YDDQESVKS 9
           |||||||||
Db       7 yddqesvks 15
```

```
RESULT  6
AAW45151
ID   AAW45151 standard; peptide; 16 AA.
XX
AC   AAW45151;
XX
DT   28-APR-1998  (first entry)
XX
DE   Human cartilage glycoprotein 39 derived peptide #52.
XX
KW   Articular cartilage; immunosuppressive therapy; antigen; HC gp 39;
KW   immunological tolerance; T-cell; human cartilage glycoprotein 39;
KW   HC gp 39; rheumatoid arthritis; epitope.
XX
OS   Synthetic.
OS   Homo sapiens.
XX
PN   WO9743008-A1.
XX
PD   13-NOV-1997.
XX
PF   22-APR-1997;  97WO-EP02051.
XX
PR   24-APR-1996;  96EP-0231106.
XX
PA   (ALKO ) AKZO NOBEL NV.
XX
PI   Boots AMH;  Verheijden GFM;
XX
DR   WPI; 1997-505775/46.
XX
PT   Peptide suitable for use in antigen specific immunosuppressive
PT   therapy - resembles or mimics epitope present on HC gp 39, so
PT   inducing systemic immunological tolerance to rheumatoid arthritis
PT   autoantigen
XX
PS   Disclosure; Page 17; 82pp; English.
XX
CC   The present sequence represents a peptide which resembles or mimics an
CC   epitope present on human cartilage glycoprotein 39 (HC gp 39), an
CC   autoantigen in rheumatoid arthritis. The invention relates to peptides
CC   consisting of 16-75 amino acid residues comprising at least one of the
CC   following 15 sequences: LVCYYISWS; FIATHIYS; IIYSEANIS; IKRIASVG;
CC   EIKSVPMT; EGHIDAWF; IVPGPPDF; YDIAFISQH; IDFISMDS; EDIPMDS;
CC   FPQQFPAP; YAVCYMEL; MLEICAPA; LAYYEICP; LSCATVHS; YLSLIAS;
CC   IAGAMVWAL; VWALDIDDF; or IDLDDFQGS. They can be used medically, in
CC   antigen specific immunosuppressive therapy, particularly the treatment
CC   of T-cell mediated destruction of articular cartilage in autoimmune
CC   diseases (e.g. rheumatoid arthritis). They can also be used to detect
CC   activated autoreactive T-cells in an individual. The peptides have a
CC   specific effect on the autoreactive T-cells, thus leaving the other
CC   components of the immune system intact, unlike the non-specific
CC   suppressive effect of immunosuppressive drugs, and do not cause toxic
CC   side effects. The peptides are predominantly recognised by autoreactive
CC   T-cells from rheumatoid arthritis patients, but rarely by those from
CC   healthy donors.
XX
SQ   Sequence  16 AA;


  Query Match             100.0%;  Score 46;  DB 2;  Length 16;
  Best Local Similarity   100.0%;  Pred. No. 0.015;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 YDDQESVKS 9
           |||| | ||
Db       1 yddqesvks 9


RESULT  7
AA002074
ID   AA002074 standard; Protein; 105 AA.
XX
AC   AA002074;
XX
DT   07-SEP-2001  (first entry)
XX
DE   Synthetic human target autoantigen sb001/8A.
XX
KW   sb001/8A; immunogenic epitope cluster; 8A;
KW   synthetic human chondrocyte glycoprotein 39/cartilage link protein;
KW   autoantigen; autoimmune disease; multiple sclerosis;
KW   rheumatoid arthritis; insulin dependent diabetes mellitus;
KW   myasthenia gravis; uveitis; autoimmune hepatitis; thyroiditis; orchitis;
KW   orchitis; idiopathic thrombocytopenic purpura; inflammatory disease;
KW   Crohn's disease; ulcerative colitis.
XX
OS   Synthetic.
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 105
FT                   /label  OTHER
FT                   /note  "other  STOP"
XX
PN   WO200131047-A2.
XX
PD   03-MAY-2001.
XX
PF   26-OCT-2000; 2000WO-IL00686.
XX
PR   27-OCT-1999;  99IL-013611.
XX
PA   (YEDA ) YEDA RES & DEV CO LTD.
XX
PI   Ben Nun A,  Kerlero De Rosbo N,  Sappier JP;
XX
DR   WPI; 2001-300515/31.
DR   N-PSDB; AAS04827.
XX
PT   Novel synthetic human target autoantigen gene useful for treating
PT   autoimmune diseases such as multiple sclerosis, insulin dependent
PT   diabetes mellitus, rheumatoid arthritis, myasthenia gravis, and uveitis
PT
XX
PS   Claim 173; Fig 41; 182pp; English.
XX
CC   The sequence represents sb001/8A (synthetic human chondrocyte
CC   glycoprotein 69/cartilage link protein) containing immunogenic epitope
CC   clusters (IEC) from chondrocyte glycoprotein 39 and cartilage link
CC   protein. The synthetic human target autoantigen genes of the invention
CC   comprise sequences coding for at least 2 IECs of autoantigen(s) related
CC   to a specific autoimmune disease. The synthetic human target autoantigen
CC   genes are useful for treating autoimmune diseases such as multiple
CC   sclerosis, insulin dependent diabetes mellitus, rheumatoid arthritis,
CC   myasthenia gravis, uveitis, autoimmune hepatitis, thyroiditis, orchitis,
CC   orchitis, idiopathic thrombocytopenic purpura, and inflammatory diseases
CC   (Crohn's disease, ulcerative colitis). The synthetic human target
CC   autoantigen genes are also useful for diagnosis and/or prediction
CC   progression of the autoimmune disease.
XX
SQ   Sequence  105 AA;


  Query Match             100.0%;  Score 46;  DB 22;  Length 105;
  Best Local Similarity   100.0%;  Pred. No. 0.14;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 YDDQESVKS 9
           |||||||||
Db      47 yddqesvks 55


RESULT  8
AA002078
ID   AA002078 standard; Protein; 22 AA.
```

```
XX
AC   AAB02079;
XX
DT   07-SEP-2001  (first entry)
XX
DE   Synthetic human multi-target autoantigen Y-RAPd.
XX
KW   Y-RAPd; immunogenic epitope cluster; IEC;
KW   autoimmune disease; multiple sclerosis; rheumatoid arthritis;
KW   insulin dependent diabetes mellitus; myasthenia gravis; uveitis;
KW   autoimmune hepatitis; thyroiditis; insulitis; orchitis;
KW   idiopathic thrombocytopenic purpura; inflammatory disease;
KW   Crohn's disease; ulcerative colitis; shMultiTAG.
XX
OS   Synthetic.
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Misc difference 279
FT                   /label= OTHER
FT                   /note= "other= STOP"
XX
PN   WO200131037-A2.
XX
PD   03-MAY-2001.
XX
PF   26-OCT-2000; 2000WO-IL00688.
XX
PR   27-OCT-1999;  99IL-0132611.
XX
PA   (YEDA ) YEDA RES & DEV CO LTD.
XX
PI   Ben-Nun A,  Kerlero De Rosbo N,  Sappler GP;
XX
DR   WPI; 2001-300515/31.
DR   N-PSDB; AAS04840.
XX
PT   Novel synthetic human target autoantigen gene useful for treating
PT   autoimmune diseases such as multiple sclerosis, insulin dependent
PT   diabetes mellitus, rheumatoid arthritis, myasthenia gravis, and uveitis
PT   -
XX
PS   Claim 188; Fig 61; 182pp; English.
XX
CC   The sequence represents synthetic human multi-target autoantigen
CC   (shMultiTAG) Y-RAPd consisting of preferred epitopes encoded by
CC   sh<Collagen>A, sh<Aggrecan>A, shHSP>A and shMBP>A. The synthetic human
CC   target autoantigen genes of the invention comprise sequences coding for
CC   at least 2 IEC's of autoantigen(s) related to a specific autoimmune
CC   disease. The synthetic human target autoantigen genes are useful for
CC   treating autoimmune diseases such as multiple sclerosis, insulin-
CC   dependent diabetes mellitus, rheumatoid arthritis, myasthenia gravis,
CC   uveitis, autoimmune hepatitis, thyroiditis, insulitis, orchitis,
CC   idiopathic thrombocytopenic purpura, and inflammatory diseases (Crohn's
CC   disease, ulcerative colitis). The synthetic human target autoantigen
CC   genes are also useful for diagnosis and/or monitoring the progression of
CC   the autoimmune disease.
XX
SQ   Sequence   281 AA;


  Query Match             100.0%;  Score 46;  DB 22;  Length 281;
  Best Local Similarity   100.0%;  Pred. No. 0.38;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 YDDQESVKS 9
           |||||||||
Db     203 yddqesvks 211
```

```
ID   AAB02079 standard; Protein; 220 AA.
XX
AC   AAB02079;
XX
DT   07-SEP-2001  (first entry)
XX
DE   Synthetic human multi-target autoantigen Y-RAPd.
XX
KW   Y-RAPd; immunogenic epitope cluster; IEC;
KW   autoimmune disease; multiple sclerosis; rheumatoid arthritis;
KW   insulin dependent diabetes mellitus; myasthenia gravis; uveitis;
KW   autoimmune hepatitis; thyroiditis; insulitis; orchitis;
KW   idiopathic thrombocytopenic purpura; inflammatory disease;
KW   Crohn's disease; ulcerative colitis; shMultiTAG.
XX
OS   Synthetic.
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Misc difference 220
FT                   /label= OTHER
FT                   /note= "other= STOP"
XX
PN   WO200131037-A2.
XX
PD   03-MAY-2001.
XX
PF   26-OCT-2000; 2000WO-IL00688.
XX
PR   27-OCT-1999;  99IL-0132611.
XX
PA   (YEDA ) YEDA RES & DEV CO LTD.
XX
PI   Ben-Nun A,  Kerlero De Rosbo N,  Sappler GP;
XX
DR   WPI; 2001-300515/31.
DR   N-PSDB; AAS04840.
XX
PT   Novel synthetic human target autoantigen gene useful for treating
PT   autoimmune diseases such as multiple sclerosis, insulin dependent
PT   diabetes mellitus, rheumatoid arthritis, myasthenia gravis, and uveitis
PT   -
XX
PS   Claim 188; Fig 61; 182pp; English.
XX
CC   The sequence represents synthetic human multi-target autoantigen
CC   (shMultiTAG) Y-RAPd consisting of preferred epitopes encoded by
CC   sh<Collagen>A, sh<Aggrecan>A, shHSP>A and shMBP>A. The synthetic human
CC   target autoantigen genes of the invention comprise sequences coding for
CC   at least 2 IEC's of autoantigen(s) related to a specific autoimmune
CC   disease. The synthetic human target autoantigen genes are useful for
CC   treating autoimmune diseases such as multiple sclerosis, insulin-
CC   dependent diabetes mellitus, rheumatoid arthritis, myasthenia gravis,
CC   uveitis, autoimmune hepatitis, thyroiditis, insulitis, orchitis,
CC   idiopathic thrombocytopenic purpura, and inflammatory diseases (Crohn's
CC   disease, ulcerative colitis). The synthetic human target autoantigen
CC   genes are also useful for diagnosis and/or monitoring the progression of
CC   the autoimmune disease.
XX
SQ   Sequence   281 AA;


  Query Match             100.0%;  Score 46;  DB 22;  Length 281;
  Best Local Similarity   100.0%;  Pred. No. 0.38;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 YDDQESVKS 9
           |||||||||
Db     203 yddqesvks 211


RESULT   9
AAB02079
```

```
RESULT  10
AAP81342
ID   AAP81342 standard; protein; 321 AA.
XX
AC   AAP81342;
XX
DT   19-OCT-1990  (first entry)
XX
DE   Polypeptide involved in protective mechanisms.
XX
KW   Immune response; cell growth.
XX
PN   JP63034898-A.
XX
PD   10-JAN-1988.
XX
PF   26-JUL-1986;   86JP-016518.
XX
PR   26-JUL-1986;   86JP-016518.
XX
PA   (DAIN ) DAINIPPON PHARM KK.
XX
DR   WPI; 1988-059819/09.
DR   P-PSDB; AAP81342.
XX
PT   New polypeptide and DNA encoding it - related to protective
PT   mechanisms such as immune response etc.
XX
PS   Disclosure; 7 pp; Japanese.
XX
CC   This polypeptide is involved in protective mechanisms such as immune
CC   response, cell growth and activation of protective functions.
XX
SQ   Sequence   321 AA;


  Query Match             100.0%;  Score 46;  DB 9;  Length 321;
  Best Local Similarity   100.0%;  Pred. No. 0.44;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 YDDQESVKS 9
           |||||||||
Db     266 yddqesvks 274


RESULT  11
AAU02076
ID   AAU02076 standard; Protein; 850 AA.
XX
AC   AAU02076;
XX
DT   07-SEP-2001  (first entry)
XX
DE   Synthetic human multi target autoantigen Y-RAPa.
XX
KW   Y-RAPa; immunogenic epitope cluster; IEC;
KW   autoantigen; shMultiTAG; RAPa; shA; shB;
KW   autoimmune disease; multiple sclerosis; rheumatoid arthritis;
KW   insulin dependent diabetes mellitus; myasthenia gravis; uveitis;
KW   autoimmune hepatitis; thyroiditis; insulitis; orchitis;
KW   idiopathic thrombocytopenic purpura; inflammatory disease;
KW   Crohn's disease; ulcerative colitis; shMultiTAG.
XX
OS   Synthetic.
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 448
FT                   /label= "HBW"
FT                   /note= "other  STOP"
XX
PN   WO200158947-A2.
XX
PD   16-AUG-2001.
XX
PF   26-OCT-2000; 2000WO-IL00688.
XX
PR   27-OCT-1999;  99US-01432611.
XX
PA   (YEDA ) YEDA RES & DEV CO LTD.
XX
PI   Ben Nun A,  Kerlero De Rosbo N,  Sappler G;
XX
DR   WPI; 2001-500515/54.
DR   N-PSDB; AAS04837.
XX
PT   Novel synthetic human target autoantigen gene useful in treating
PT   autoimmune diseases such as multiple sclerosis, insulin dependent
PT   diabetes mellitus, rheumatoid arthritis, myasthenia gravis and uveitis
PT
XX
PS   Claim 179; Fig 55; 182pp; English.
XX
CC   The sequence represents synthetic human multi target autoantigen
CC   (shMultiTAG) Y-RAPa consisting of regions encoding shCellTag/RA
CC   shA/retroCam/RA and shB/LF/RA. The synthetic human target autoantigen(s)
CC   of the invention comprise sequences coding for at least 2 IEC(s) of
CC   autoantigen(s) related to a specific autoimmune disease. The synthetic
CC   human target autoantigen genes are useful for treating autoimmune
CC   diseases such as multiple sclerosis, insulin dependent diabetes
CC   mellitus, rheumatoid arthritis, myasthenia gravis, uveitis, autoimmune
CC   hepatitis, thyroiditis, insulitis, orchitis, idiopathic thrombocytopenic
CC   purpura, and inflammatory diseases (Crohn's disease, ulcerative colitis).
CC   The synthetic human target autoantigen genes are also useful for
CC   diagnosis and/or monitoring the progression of the autoimmune disease
XX
SQ   Sequence   850 AA;


  Query Match             100.0%;  Score 46;  DB 22;  Length 850;
  Best Local Similarity   100.0%;  Pred. No. 0.49;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 YDDQESVKS 9
           |||||||||
Db     292 yddqesvks 300


RESULT  12
AAU02077
ID   AAU02077 standard; Protein; 459 AA.
XX
AC   AAU02077;
XX
DT   07-SEP-2001  (first entry)
XX
DE   Synthetic multi target autoantigen Y-RAPb.
XX
KW   Y-RAPa; immunogenic epitope cluster; IEC;
KW   autoantigen; shMultiTAG; RAPa; shA; shB;
KW   autoimmune disease; multiple sclerosis; rheumatoid arthritis;
KW   insulin dependent diabetes mellitus; myasthenia gravis; uveitis;
KW   autoimmune hepatitis; thyroiditis; insulitis; orchitis;
KW   idiopathic thrombocytopenic purpura; inflammatory disease;
KW   Crohn's disease; ulcerative colitis; shMultiTAG.
XX
OS   Synthetic.
OS   Homo sapiens.
OS   Escherichia coli.
OS   Mycobacterium tuberculosis.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 455
FT                   /label= "HBW"
FT                   /note= "other  STOP"
```

```
FT   Misc-difference 45..
FT                   /label= OTHER
FT                   /note= "other - SIGP"
XX
PN   WO200131337-A2.
XX
PD   03-MAY-2001.
XX
PF   27-OCT-2000; 2000WO-IL00688.
XX
PR   27-OCT-1999;  1999IL-0132611.
XX
PA   (YEDA ) YEDA RES & DEV CO LTD.
XX
PI   Ben Nun A;  Kerner Or;  De Rosbo N;  Sappler GP;
XX
DR   WPI; 2001-300519/31
DR   N-PSDB; AAX81838.
XX
PT   Novel synthetic human target autoantigen gene useful for treating
PT   autoimmune diseases such as multiple sclerosis, insulin-dependent
PT   diabetes mellitus, rheumatoid arthritis, myasthenia gravis, and uveitis
PT
XX
PS   Claim 182; Fig 17; 282pp; English.
XX
CC   The sequence represents synthetic human multi target autoantigen
CC   (SHuMuTiAG) Y (ABs consisting of regions encoding shCollagen/kA,
CC   shAggrecan/kA, shS-S/kA or shHSP65/kA. The synthetic human target
CC   autoantigen genes of the invention comprise sequences coding for at least
CC   2 IECs of autoantigen(s) related to a specific autoimmune disease. The
CC   synthetic human target autoantigen genes are useful for treating
CC   autoimmune diseases such as multiple sclerosis, insulin-dependent
CC   diabetes mellitus, rheumatoid arthritis, myasthenia gravis, uveitis,
CC   autoimmune hepatitis, thyroiditis, insulitis, orchitis, idiopathic
CC   thrombocytopenic purpura and inflammatory diseases (Crohn's disease,
CC   ulcerative colitis). The synthetic human target autoantigen genes are
CC   also useful for diagnosis and/or monitoring the progression of the
CC   autoimmune disease
XX
SQ   Sequence   459 AA;
```

```
  Query Match             10.0%;  Score 46;  DB 22;  Length 459;
  Best Local Similarity   10.0%;  Pred. No. 0.67;
  Matches    9;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 YDDQESVKS 9
           |:|  ||||
Db     292 yddqesvks 300
```

```
RESULT  13
AAW26751
ID   AAW26751 standard; Protein; 383 AA.
XX
AC   AAW26751;
XX
DT   11-MAY-1998  (first entry)
XX
DE   Bovine whey protein.
XX
KW   Bovine whey protein; human cartilage glycoprotein 39; HC gp-39;
KW   autoantigen; antigen; autoimmune disease; rheumatoid arthritis;
KW   inflammation; arthritogenic protein; immunotherapy; therapy.
XX
OS   Bos taurus.
XX
PN   WO9741147-A1.
XX
PD   30-OCT-1997.
XX
PF   18-APR-1997;  97WO-EP01955.
XX
PR   18-APR-1996;  96US-0634493.
XX
PA   (ALKU ) AKZO NOBEL NV.
XX
PI   Boots AMH;  Bos ES;  Verheijden GFM;
XX
DR   WPI; 1997-535833/49.
DR   N-PSDB; AAT99452.
XX
PT   Autoantigen proteins homologous to human cartilage glycoprotein 39 -
PT   induce arthritis and provide antigen specific treatment of articular
PT   cartilage destruction in autoimmune diseases e.g. rheumatoid
PT   arthritis
XX
PS   Claim 11; Page 23-25; 35pp; English.
XX
CC   This bovine 39 kDa whey protein sequence was deduced from cDNA
CC   clones (see AAT99452) isolated from a cattle mammary gland cDNA
CC   library using human cartilage glycoprotein 39 (HC gp 39) cDNA as
CC   probe. Bovine whey protein was shown to be arthritogenic, inducing
CC   arthritis in animals in the same way as described for HC gp 39 (see
CC   AAW26750). Arthritogenic proteins such as bovine whey protein, and
CC   which display at least 50% amino acid homology to HC gp-39, are
CC   very suitable for inducing systemic tolerance of the immune system
CC   to homologous autoantigens and can be used to delay and/or suppress
CC   arthritic development in mammals. They can induce specific T cell
CC   tolerance to HC gp-39 in patients suffering from T-cell mediated
CC   cartilage destruction, such as rheumatoid arthritis. Arthritogenic
CC   proteins are also suitable to induce arthritis in animals,
CC   preferably mice, e.g. for use in drug screening.
XX
SQ   Sequence   383 AA;
```

```
  Query Match             93.5%;  Score 43;  DB 18;  Length 383;
  Best Local Similarity   88.9%;  Pred. No. 2.1;
  Matches    8;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 YDDQESVKS 9
           |||||||||:
Db     328 yddqesvkn 336
```

```
RESULT  14
AAB03442
ID   AAB03442 standard; Protein; 383 AA.
XX
AC   AAB03442;
XX
DT   03-JAN-2001  (first entry)
XX
DE   Gp38k protein sequence.
XX
KW   Gp38k; chemoattractant; cell migration; wound healing; angiogenesis;
KW   cancer; vascular trauma; vascular disease; atherosclerosis; restenosis;
XX
OS   Unidentified.
XX
FH   Key             Location/Qualifiers
FT   Peptide         1..24
FT                   /label= signal_peptide
FT   Region          30..37
FT                   /note= "peptide antibody"
FT   Modified-site   59..61
FT                   /label= glycosylation_site
FT   Binding-site    68..75
FT                   /label= leucine_zipper
FT   Binding-site    82..89
FT                   /label= leucine_zipper
FT   Active-site     131..136
FT                   /label= chitanase_active_site
FT   Binding-site    143..146
```

```
FT                   /label= heparin_binding_site
FT   Binding site    147..154
FT                   /label= hyaluronic_acid_binding_site
FT   Binding site    262..270
FT                   /label= hyaluronic_acid_binding_site
FT   Binding site    278..281
FT                   /label= glycosaminoglycan_binding site
FT   Region          454..457
FT                   /note= "acidic region"
FT   Binding site    468..476
FT                   /label= hyaluronic acid_binding site
XX
PN   WO200044776-A1.
XX
PD   15-JUN-2000.
XX
PF   10-DEC-1999;    99WO-US29262.
XX
PR   11-DEC-1998;    98US-0111856.
XX
PA   (UYNY-) UNIV NEW YORK STATE RES FOUND.
XX
PI   Millis AJT;
XX
DR   WPI; 2000-417457/37.
XX
PT   Clusterin and apo8K-related peptide capable of altering cell migration
PT   useful for treating atherosclerosis, cancer and stenosis following
PT   vascular trauma or disease
XX
PS   Disclosure; Fig 2; 43pp; English.
XX
CC   The present sequence is the protein sequence of apo8K, apo8K, a
CC   chemoattractant, is essential for the migration of vascular smooth muscle
CC   cells (VSMC). The gene and protein can, therefore, be used to promote
CC   wound healing, angiogenesis and vasculogenesis, in the treatment of
CC   stenosis following vascular trauma or disease and to treat
CC   atherosclerosis, and antisense sequences can be used to treat cancer, as
CC   angiogenesis is vital for tumour survival.
XX
SQ   Sequence   484 AA;


  Query Match             93.5%;  Score 44;  DB 21;  Length 484;
  Best Local Similarity   88.9%;  Pred. No. 2.1;
  Matches    8;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 YDDQESVKS 9
           ||||:|||:
Db     428 ydddqesvkn 436


RESULT  15
ABR55896
ID   ABR55896 standard; Peptide; 14 AA.
XX
AC   ABR55896;
XX
DT   15-FEB-2002  (first entry)
XX
DE   Vascular dementia associated protein isoform (VPI) #6.
XX
KW   Vascular Dementia; VD; VD-associated protein isoform; VPI; screening;
KW   diagnosis; prognosis; gene therapy.
XX
OS   Homo sapiens.
XX
PN   WO200167526-A1.
XX
PD   20-SEP-2001.
XX
PF   14-MAR-2000; 2001WO-GB01106.
XX
PR   15-MAR-2000; 2000GB-0006285.
PR   24-NOV-2000; 2000GB-0028734.
PR   28-NOV-2000; 2000US-0724791.
XX
PA   (OXFO-) OXFORD GLYCOSCIENCES UK LTD.
XX
PI   Herath HMAC;  Parekh RB;  Rohlff C;
XX
DR   WPI; 2001-557963/62.
XX
PT   Screening, diagnosis or prognosis of vascular dementia (VD), useful for
PT   determining stage of VD and monitoring the effect of VD therapy,
PT   comprises analysing body fluid by 2 dimensional electrophoresis
PT   features correlated with VD
XX
PS   Claim 6; Page 5; 151pp; English.
XX
CC   The invention relates to screening, diagnosis or prognosis of Vascular
CC   Dementia (VD) in a subject comprising analysing body fluid from the
CC   subject by 2 dimensional (2-D) electrophoresis to generate a 2-D array of
CC   features containing at least one chosen feature whose relative abundance
CC   correlates with the presence, absence, stage or severity of VD and
CC   predicts the onset or course of VD, especially detecting in a sample of
CC   cerebrospinal fluid (CSF) from the subject one of 22 VD associated
CC   protein isoforms (VPIs) (Abb55891-Abb55912), as fully defined in the
CC   specification. Detecting VD associated features and VPI is useful for the
CC   screening, diagnosis or prognosis of VD, for determining the stage or
CC   severity of VD, for identifying a subject at risk of VD or for
CC   monitoring the effect of therapy administered to a subject having VD.
CC   Nucleic acids encoding a VPI or inhibiting the function of a VPI are
CC   useful for the treatment of VD and for gene therapy.
XX
SQ   Sequence   14 AA;


  Query Match             93.5%;  Score 44;  DB 22;  Length 14;
  Best Local Similarity   88.9%;  Pred. No. 0.0078;
  Matches    8;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 YDDQESVKS 9
           ||||:|||:
Db       7 yddqesvkn 15


Search completed: August 26, 2002, 06:13:24
Job time: 266 sec
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 26, 2002, 06:23:26 ; Search Time 46.22 Seconds
(without alignments)
4.756 Million cell updates/sec

Title: US-09-744-282-3
Perfect score: 46
Sequence: 1 YDDQESVKS 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 237428 seqs, 244559941 residues

Total number of hits satisfying chosen parameters: 237428

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : issued_patents_AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	46	100.0	9	1	US-08-619-645-3	Sequence 3, Appli
2	46	100.0	9	2	US-08-644-494-3	Sequence 3, Appli
3	46	100.0	13	1	US-08-619-645-7	Sequence 7, Appli
4	46	100.0	13	2	US-08-644-494-7	Sequence 7, Appli
5	46	100.0	16	4	US-09-171-705-51	Sequence 51, Appl
6	46	100.0	16	4	US-09-171-705-52	Sequence 52, Appl
7	32	69.6	468	1	US-08-204-656B-2	Sequence 2, Appli
8	32	69.6	468	1	US-08-204-656B-4	Sequence 4, Appli
9	32	69.6	468	1	US-08-204-656B-6	Sequence 6, Appli
10	32	69.6	468	1	US-08-204-656B-8	Sequence 8, Appli
11	32	69.6	468	1	US-08-470-702-6	Sequence 6, Appli
12	32	69.6	468	1	US-08-470-702-7	Sequence 7, Appli
13	32	69.6	468	1	US-08-470-702-8	Sequence 8, Appli
14	32	69.6	468	1	US-08-470-702-9	Sequence 9, Appli
15	32	69.6	468	1	US-08-467-831-6	Sequence 6, Appli
16	32	69.6	468	1	US-08-467-831-7	Sequence 7, Appli
17	32	69.6	468	1	US-08-467-831-8	Sequence 8, Appli
18	32	69.6	468	1	US-08-467-831-9	Sequence 9, Appli
19	31	67.4	188	2	US-08-933-750C-39	Sequence 39, Appl
20	31	67.4	188	4	US-09-234-614-39	Sequence 39, Appl
21	31	67.4	471	2	US-08-591-629-2	Sequence 2, Appli
22	30	65.2	467	3	US-08-467-023-95	Sequence 95, Appl
23	30	65.2	469	1	US-08-074-121-6	Sequence 6, Appli
24	30	65.2	469	5	PCT-US94-06447-6	Sequence 6, Appli
25	30	65.2	489	1	US-07-903-103-4	Sequence 4, Appli
26	30	65.2	489	1	US-08-044-619A-4	Sequence 4, Appli
27	30	65.2	489	1	US-08-283-911-4	Sequence 4, Appli
28	30	65.2	489	1	US-08-245-500A-5	Sequence 5, Appl
29	30	65.2	489	1	US-08-390-546-5	Sequence 5, Appl
30	30	65.2	489	1	US-08-390-479A-5	Sequence 5, Appl
31	30	65.2	489	1	US-08-557-393-5	Sequence 5, Appl
32	30	65.2	489	1	US-08-390-516C-5	Sequence 5, Appli
33	30	65.2	489	1	US-08-390-517A-5	Sequence 5, Appli
34	30	65.2	489	1	US-08-390-515A-5	Sequence 5, Appli
35	30	65.2	489	2	US-08-801-718-5	Sequence 5, Appli
36	30	65.2	514	2	US-09-122-230-7	Sequence 7, Appli
37	30	65.2	1013	4	US-08-960-886-2	Sequence 2, Appli
38	30	65.2	1014	4	US-09-078-347A-3	Sequence 3, Appli
39	29	63.0	28	2	US-08-520-535-26	Sequence 26, Appl
40	29	63.0	28	2	US-09-079-432-26	Sequence 26, Appl
41	29	63.0	60	6	5204445-5	Patent No. 5204445
42	29	63.0	64	6	5204445-6	Patent No. 5204445
43	29	63.0	68	6	5204445-7	Patent No. 5204445
44	29	63.0	251	5	PCT-US96-01514-59	Sequence 59, Appl
45	29	63.0	674	3	US-08-893-852A-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1
US-08-619-645-3
; Sequence 3, Application US/08619645
; Patent No. 5736507
; GENERAL INFORMATION:
; APPLICANT: Boots, Anna M.H.
; APPLICANT: Verheijden, Gijsbertus F.M.
; TITLE OF INVENTION: N- [illegible] peptides derived from
; TITLE OF INVENTION: autoantigen for use in immunotherapy of autoimmune
; TITLE OF INVENTION: disease
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Akzo Nobel Patent Department
; STREET: 1300 Piccard Drive, Suite 206
; CITY: Rockville
; STATE: Maryland
; COUNTRY: U.S.A.
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25(EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/619,645
; FILING DATE: 25-MAR-1996
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Gormley, Mary E.
; REGISTRATION NUMBER: 34,409
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 947-4433
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-619-645-3

Query Match 100.0%; Score 46; DB 1; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.7e-05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YDDQESVKS 9
|||||||||
Db 1 YDDQESVKS 9

```
RESULT   2
US-08-644-493-3
; Sequence 3, Application US/08644493
; Patent No. 5844449
;  GENERAL INFORMATION:
;     APPLICANT:  A.M.H. Boots
;     APPLICANT:  G.F.M. Verheijden
;     APPLICANT:  F.S. Bos
;     TITLE OF INVENTION:  No. 5844449el Peptides derived from autoantigen for use
;     TITLE OF INVENTION:  in Immunotherapy of Autoimmune Diseases
;     NUMBER OF SEQUENCES:  10
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Akzo No. 5844449el Patent Department
;       STREET:  1300 Piccard Drive, Suite 206
;       CITY:  Rockville
;       STATE:  Maryland
;       COUNTRY:  USA
;       ZIP:  20850
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/644,493
;       FILING DATE:  18-APR-1996
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/619,645
;       FILING DATE:  25-MAR-1996
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  PCT/EP95/04201
;       FILING DATE:  25-OCT-1995
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  NL 9420l287
;       FILING DATE:  27-OCT-1994
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  NL 95200860
;       FILING DATE:  07-APR-1995
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (30l) 258-5200
;       TELEFAX:  (30l) 977-0847
;   INFORMATION FOR SEQ ID NO:  3:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  9 amino acids
;       TYPE:  amino acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  peptide
US-08-644-493-3


  Query Match             100.0%;  Score 46;  DB 2;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 1.7e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 YDDQESVKS 9
           |||||||||
Db       1 YDDQESVKS 9


RESULT   3
US-08-619-645-7
; Sequence 7, Application US/08619645
; Patent No. 5736507
;  GENERAL INFORMATION:
;     APPLICANT:  Boots, Anna M.H.
;     APPLICANT:  Verheijden, Gijsbertus F.M.
;     TITLE OF INVENTION:  No. 5736507el peptides derived from
;     TITLE OF INVENTION:  autoantigens for use in Immunotherapy of Autoimmune
;     TITLE OF INVENTION:  disease
;     NUMBER OF SEQUENCES:  9
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Akzo No. 5736507el Patent Department
;       STREET:  1300 Piccard Drive, Suite 206
;       CITY:  Rockville
;       STATE:  Maryland
;       COUNTRY:  U.S.A.
;       ZIP:  20850
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25(EPO)
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/619,645
;       FILING DATE:  25-MAR-1996
;       CLASSIFICATION:  514
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Gormley, Mary E.
;       REGISTRATION NUMBER:  34,409
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (30l) 947-4433
;   INFORMATION FOR SEQ ID NO:  7:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  13 amino acids
;       TYPE:  amino acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  peptide
US-08-619-645-7


  Query Match             100.0%;  Score 46;  DB 1;  Length 13;
  Best Local Similarity   100.0%;  Pred. No. 0.00024;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 YDDQESVKS 9
           |||||||||
Db       3 YDDQESVKS 11


RESULT   4
US-08-644-493-7
; Sequence 7, Application US/08644493
; Patent No. 5844449
;  GENERAL INFORMATION:
;     APPLICANT:  A.M.H. Boots
;     APPLICANT:  G.F.M. Verheijden
;     APPLICANT:  F.S. Bos
;     TITLE OF INVENTION:  No. 5844449el Peptides derived from autoantigen for use
;     TITLE OF INVENTION:  in Immunotherapy of Autoimmune Diseases
;     NUMBER OF SEQUENCES:  10
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Akzo No. 5844449el Patent Department
;       STREET:  1300 Piccard Drive, Suite 206
;       CITY:  Rockville
;       STATE:  Maryland
;       COUNTRY:  USA
;       ZIP:  20850
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/644,493
;       FILING DATE:  18-APR-1996
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/619,645
;       FILING DATE:  25-MAR-1996
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  PCT/EP95/04201
;       FILING DATE:  25-OCT-1995
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  NL 9420l287
```

```
;       FILING DATE:  27-OCT-1994
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  NL 9420038860
;       FILING DATE:  07-APR-1995
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  ( 31) 258 5200
;       TELEFAX:  (301) 907-0847
;   INFORMATION FOR SEQ ID NO:  7:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  13 amino acids
;       TYPE:  amino acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  peptide
US-08-634-493-7


  Query Match             100.0%;  Score 46;  DB 2;  Length 13;
  Best Local Similarity   100.0%;  Pred. No. 0.0023;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 YDDQESVKS 9
            |||||||||
Db       3 YDDQESVKS 11


RESULT   5
US-09-171-705-51
; Sequence 51, Application US/09171705
; Patent No. 6184204
; GENERAL INFORMATION:
;  APPLICANT: BOOTS, ANNA M.H.
;  APPLICANT: VERHEIJDEN, GILBERTUS F.M.
;  TITLE OF INVENTION: NOVEL PEPTIDES SUITABLE FOR USE IN ANTIGEN SPECIFIC
;  TITLE OF INVENTION: IMMUNOSUPPRESSIVE THERAPY
;  FILE REFERENCE: O/96198 US
;  CURRENT APPLICATION NUMBER: US/09/171,705
;  CURRENT FILING DATE: 1999-02-09
;  NUMBER OF SEQ ID NOS: 78
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 51
;   LENGTH: 16
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: DERIVED FROM
;   OTHER INFORMATION: SEQUENCE OF HUMAN CARTILAGE (HC) -39 PROTEIN
US-09-171-705-51


  Query Match             100.0%;  Score 46;  DB 4;  Length 16;
  Best Local Similarity   100.0%;  Pred. No. 0.0028;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 YDDQESVKS 9
            |||||||||
Db       7 YDDQESVKS 15


RESULT   6
US-09-171-705-52
; Sequence 52, Application US/09171705
; Patent No. 6184204
; GENERAL INFORMATION:
;  APPLICANT: BOOTS, ANNA M.H.
;  APPLICANT: VERHEIJDEN, GILBERTUS F.M.
;  TITLE OF INVENTION: NOVEL PEPTIDES SUITABLE FOR USE IN ANTIGEN SPECIFIC
;  TITLE OF INVENTION: IMMUNOSUPPRESSIVE THERAPY
;  FILE REFERENCE: O/96198 US
;  CURRENT APPLICATION NUMBER: US/09/171,705
;  CURRENT FILING DATE: 1999-02-09
;  NUMBER OF SEQ ID NOS: 78
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 52
;   LENGTH: 16
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: DERIVED FROM
;   OTHER INFORMATION: SEQUENCE OF HUMAN CARTILAGE (HC) -39 PROTEIN
US-09-171-705-52


  Query Match             100.0%;  Score 46;  DB 4;  Length 16;
  Best Local Similarity   100.0%;  Pred. No. 0.0028;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 YDDQESVKS 9
            |||||||||
Db       1 YDDQESVKS 9


RESULT   7
US-08-204-656B-2
; Sequence 2, Application US/08204656B
; Patent No. 5538882
;  GENERAL INFORMATION:
;    APPLICANT:  Matsui, Ikuo
;    APPLICANT:  Ishikawa, Kazuhiko
;    APPLICANT:  Miyairi, Sachio
;    APPLICANT:  Honda, Koichi
;    TITLE OF INVENTION:  Variant-Type Carbohydrate Hydrolase,
;    TITLE OF INVENTION:  Variant Gene Of The Enzyme And Method For Producing
;    TITLE OF INVENTION:  Oligosaccharide Using The Enzyme
;    NUMBER OF SEQUENCES:  18
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Birch, Stewart, Kolasch & Birch, LLP
;      STREET:  8110 Gatehouse Road, Suite 500 East
;      CITY:  Falls Church
;      STATE:  Virginia
;      COUNTRY:  U.S.A.
;      ZIP:  22042
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/204,656B
;      FILING DATE:  02-MAR-1994
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Weiner, Marc S.
;      REGISTRATION NUMBER:  32,181
;      REFERENCE/DOCKET NUMBER:  234 252P
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (703) 205-8000
;      TELEFAX:  (703) 205-8050
;      TELEX:  248345
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  468 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-204-656B-2


  Query Match             69.6%;  Score 32;  DB 1;  Length 468;
  Best Local Similarity   66.7%;  Pred. No. 73;
  Matches    6;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 YDDQESVKS 9
            ||||  |:|
Db     156 YDDQAQVQS 164
```

```
RESULT  8
US-08-204-656B-4
; Sequence 4, Application US/08204656B
; Patent No. 5548888,
;  GENERAL INFORMATION:
;     APPLICANT:  Matsui, Ikuo
;     APPLICANT:  Ishikawa, Kazuhiko
;     APPLICANT:  Miyairi, Sachio
;     APPLICANT:  Honda, Koichi
;     TITLE OF INVENTION:  Variant Type Carbohydrate Hydrolase,
;     TITLE OF INVENTION:  Variant Gene Of The Enzyme And Method For Producing
;     TITLE OF INVENTION:  Oligosaccharide Using The Enzyme
;     NUMBER OF SEQUENCES:  18
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Birch, Stewart, Kolasch & Birch, LLP
;       STREET:  8110 Gatehouse Road, Suite 500 East
;       CITY:  Falls Church
;       STATE:  Virginia
;       COUNTRY:  U.S.A.
;       ZIP:  22042
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.30
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/204,656B
;       FILING DATE:  02-MAR-1994
;       CLASSIFICATION:  435
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Weiner, Marc S.
;       REGISTRATION NUMBER:  32,181
;       REFERENCE/DOCKET NUMBER:  234 252P
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (703) 205-8000
;       TELEFAX:  (703) 205-8050
;       TELEX:  248345
;   INFORMATION FOR SEQ ID NO:  4:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  468 amino acids
;       TYPE:  amino acid
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  protein
US-08-204-656B-4


  Query Match             69.6%;  Score 32;  DB 1;  Length 468;
  Best Local Similarity   66.7%;  Prod. No. 73;
  Matches    6;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 YIDQESVKS 9
            |||:  |:|
Db     156 YIDQAQVQS 164


RESULT  9
US-08-204-656B-6
; Sequence 6, Application US/08204656B
; Patent No. 5548882
;  GENERAL INFORMATION:
;     APPLICANT:  Matsui, Ikuo
;     APPLICANT:  Ishikawa, Kazuhiko
;     APPLICANT:  Miyairi, Sachio
;     APPLICANT:  Honda, Koichi
;     TITLE OF INVENTION:  Variant Type Carbohydrate Hydrolase,
;     TITLE OF INVENTION:  Variant Gene Of The Enzyme And Method For Producing
;     TITLE OF INVENTION:  Oligosaccharide Using The Enzyme
;     NUMBER OF SEQUENCES:  18
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Birch, Stewart, Kolasch & Birch, LLP
;       STREET:  8110 Gatehouse Road, Suite 500 East
;       CITY:  Falls Church
;       STATE:  Virginia
;       COUNTRY:  U.S.A.
;       ZIP:  22042
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.30
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/204,656B
;       FILING DATE:  02-MAR-1994
;       CLASSIFICATION:  435
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Weiner, Marc S.
;       REGISTRATION NUMBER:  32,181
;       REFERENCE/DOCKET NUMBER:  234 252P
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (703) 205-8000
;       TELEFAX:  (703) 205-8050
;       TELEX:  248345
;   INFORMATION FOR SEQ ID NO:  6:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  468 amino acids
;       TYPE:  amino acid
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  protein
US-08-204-656B-6


  Query Match             69.6%;  Score 32;  DB 1;  Length 468;
  Best Local Similarity   66.7%;  Prod. No. 73;
  Matches    6;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 YIDQESVKS 9
            |||:  |:|
Db     156 YIDQAQVQS 164


RESULT  10
US-08-204-656B-8
; Sequence 8, Application US/08204656B
; Patent No. 5548882
;  GENERAL INFORMATION:
;     APPLICANT:  Matsui, Ikuo
;     APPLICANT:  Ishikawa, Kazuhiko
;     APPLICANT:  Miyairi, Sachio
;     APPLICANT:  Honda, Koichi
;     TITLE OF INVENTION:  Variant Type Carbohydrate Hydrolase,
;     TITLE OF INVENTION:  Variant Gene Of The Enzyme And Method For Producing
;     TITLE OF INVENTION:  Oligosaccharide Using The Enzyme
;     NUMBER OF SEQUENCES:  18
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Birch, Stewart, Kolasch & Birch, LLP
;       STREET:  8110 Gatehouse Road, Suite 500 East
;       CITY:  Falls Church
;       STATE:  Virginia
;       COUNTRY:  U.S.A.
;       ZIP:  22042
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.30
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/204,656B
;       FILING DATE:  02-MAR-1994
;       CLASSIFICATION:  435
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Weiner, Marc S.
;       REGISTRATION NUMBER:  32,181
;       REFERENCE/DOCKET NUMBER:  234 252P
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (703) 205-8000
```

```
;       TELEFAX:  (703) 205-8050
;       TELEX:  248345
;   INFORMATION FOR SEQ ID NO:  8:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  468 amino acids
;       TYPE:  amino acid
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  protein
US-08-204,656B-8


  Query Match             69.6%;  Score 32;  DB 1;  Length 468;
  Best Local Similarity   66.7%;  Pred. No. 73;
  Matches    6;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 YDDQESVKS 9
            |||| |:|
Db     156 YDDQAQVQS 164


RESULT  12
US-08-470-702-7
; Sequence 7, Application US/08470702
; Patent No. 5631149
;  GENERAL INFORMATION:
;     APPLICANT:  MATSUI, IKUO
;     APPLICANT:  ISHIKAWA, KAZUHIKO
;     APPLICANT:  MIYAIRI, SACHIO
;     APPLICANT:  HONDA, KOICHI
;     TITLE OF INVENTION:  VARIANT-TYPE CARBOHYDRATE HYDROLASE,
;     TITLE OF INVENTION:  VARIANT GENE OF THE ENZYME AND METHOD FOR PRODUCING
;     TITLE OF INVENTION:  OLIGOSACCHARIDE USING THE ENZYME
;     NUMBER OF SEQUENCES:  17
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  BIRCH, STEWART, KOLASCH & BIRCH
;       STREET:  8110 Gatehouse Road, Suite 500 East
;       CITY:  Falls Church
;       STATE:  Virginia
;       COUNTRY:  U.S.A.
;       ZIP:  22042
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC DOS/MS DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/470,702
;       FILING DATE:  06-JUN-1995
;       CLASSIFICATION:  435
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/204,656
;       FILING DATE:  02-MAR-1994
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  WEINER, MARC S.
;       REGISTRATION NUMBER:  32,181
;       REFERENCE/DOCKET NUMBER:  234-252P
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (703) 205-8000
;       TELEFAX:  (703) 205-8050
;       TELEX:  248345
;   INFORMATION FOR SEQ ID NO:  7:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  468 amino acids
;       TYPE:  amino acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  protein
;     HYPOTHETICAL:  NO
;     ANTI-SENSE:  NO
US-08-470-702-7


  Query Match             69.6%;  Score 32;  DB 1;  Length 468;
  Best Local Similarity   66.7%;  Pred. No. 73;
  Matches    6;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 YDDQESVKS 9
            |||| |:|
Db     156 YDDQAQVQS 164


RESULT  13
US-08-470-702-8
; Sequence 8, Application US/08470702
; Patent No. 5631149
```

```
;  GENERAL INFORMATION:
;    APPLICANT:  MATSUI, IKUO
;    APPLICANT:  ISHIKAWA, KAZUHIKO
;    APPLICANT:  MIYAIRI, SACHIO
;    APPLICANT:  HONDA, KOICHI
;    TITLE OF INVENTION:  VARIANT TYPE CARBOHYDRATE HYDROLASE,
;    TITLE OF INVENTION:  VARIANT GENE OF THE ENZYME AND METHOD FOR ITS SEQUENCING
;    TITLE OF INVENTION:  OLIGOSACCHARIDE USING THE ENZYME
;    NUMBER OF SEQUENCES:  17
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  BIRCH, STEWART, KOLASCH & BIRCH
;      STREET:  8110 Gatehouse Road, Suite 500 East
;      CITY:  Falls Church
;      STATE:  Virginia
;      COUNTRY:  U.S.A.
;      ZIP:  22042
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/470,702
;      FILING DATE:  06-JUN-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/204,656
;      FILING DATE:  02-MAR-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  WEINER, MARC S.
;      REGISTRATION NUMBER:  32,181
;      REFERENCE/DOCKET NUMBER:  234-252P
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (703) 205-8000
;      TELEFAX:  (703) 205-8050
;      TELEX:  248345
;  INFORMATION FOR SEQ ID NO:  8:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  468 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
US-08-470-702-8


  Query Match             69.6%;  Score 32;  DB 1;  Length 468;
  Best Local Similarity   66.7%;  Pred. No. 73;
  Matches    6;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 YDDQESVKS 9
           |||| |:|
Db     156 YDDQAQVQS 164


RESULT  14
US-08-470-702-9
; Sequence 9, Application US/08470702
; Patent No. 5631149
;  GENERAL INFORMATION:
;    APPLICANT:  MATSUI, IKUO
;    APPLICANT:  ISHIKAWA, KAZUHIKO
;    APPLICANT:  MIYAIRI, SACHIO
;    APPLICANT:  HONDA, KOICHI
;    TITLE OF INVENTION:  VARIANT TYPE CARBOHYDRATE HYDROLASE,
;    TITLE OF INVENTION:  VARIANT GENE OF THE ENZYME AND METHOD FOR ITS SEQUENCING
;    TITLE OF INVENTION:  OLIGOSACCHARIDE USING THE ENZYME
;    NUMBER OF SEQUENCES:  17
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  BIRCH, STEWART, KOLASCH & BIRCH
;      STREET:  8110 Gatehouse Road, Suite 500 East
;      CITY:  Falls Church
;      STATE:  Virginia
;      COUNTRY:  U.S.A.
;      ZIP:  22042
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/470,702
;      FILING DATE:  06-JUN-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/204,656
;      FILING DATE:  02-MAR-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  WEINER, MARC S.
;      REGISTRATION NUMBER:  32,181
;      REFERENCE/DOCKET NUMBER:  234-252P
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (703) 205-8000
;      TELEFAX:  (703) 205-8050
;      TELEX:  248345
;  INFORMATION FOR SEQ ID NO:  9
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  468 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
US-08-470-702-9


  Query Match             69.6%;  Score 32;  DB 1;  Length 468;
  Best Local Similarity   66.7%;  Pred. No. 73;
  Matches    6;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 YDDQESVKS 9
           |||| |:|
Db     156 YDDQAQVQS 164


RESULT  15
US-08-467-831-6
; Sequence 6, Application US/08467831
; Patent No. 5635378
;  GENERAL INFORMATION:
;    APPLICANT:  MATSUI, IKUO
;    APPLICANT:  ISHIKAWA, KAZUHIKO
;    APPLICANT:  MIYAIRI, SACHIO
;    APPLICANT:  HONDA, KOICHI
;    TITLE OF INVENTION:  VARIANT TYPE CARBOHYDRATE HYDROLASE,
;    TITLE OF INVENTION:  VARIANT GENE OF THE ENZYME AND METHOD FOR ITS SEQUENCING
;    TITLE OF INVENTION:  OLIGOSACCHARIDE USING THE ENZYME
;    NUMBER OF SEQUENCES:  17
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  BIRCH, STEWART, KOLASCH & BIRCH
;      STREET:  8110 Gatehouse Road, Suite 500 East
;      CITY:  Falls Church
;      STATE:  Virginia
;      COUNTRY:  U.S.A.
;      ZIP:  22042
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/467,831
;      FILING DATE:  06-JUN-1995
```

```
;       CLASSIFICATION:  445
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/204,656
;       FILING DATE:  01-MAR-1994
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  WEINER, MARC S.
;       REGISTRATION NUMBER:  32,181
;       REFERENCE/DOCKET NUMBER:  234.2720
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (703) 205-8000
;       TELEFAX:  (703) 205-8050
;       TELEX:  248345
;   INFORMATION FOR SEQ ID NO:  6:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  468 amino acids
;       TYPE:  amino acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  protein
;     HYPOTHETICAL:  NO
;     ANTI-SENSE:  NO
US-08-467-841-6


  Query Match             69.6%;  Score 42;  DB 1;  Length 468;
  Best Local Similarity   65.7%;  Pred. No. 75;
  Matches   6;  Conservative   1;  Mismatches   2;  Indels   0;  Gaps   0;

Qy       1 YDDQESVKS 9
             |||  | |
Db     159 YDDQAIVQS 167



Search completed: August 26, 2002, 05:51:27
Job time : 808 secs
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 26, 2002, 06:10:47 ; Search time 8.5 Seconds
(without alignments)
10,743 Million cell updates/sec

Title: US-09-744-282-4
Perfect score: 43
Sequence: 1 FSKIASNIQ 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283438 seqs, 96809634 residues

Total number of hits satisfying chosen parameters: 283438

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 71:*
1: pir1 *
2: pir2 *
3: pir3 *
4: pir4 *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	43	100.0	383	2	S51427	heparin-binding glycoprotein 88K - pig
2	43	100.0	383	2	A49562	cartilage glycopro
3	36	83.7	405	2	S61551	breast regressing
4	34	79.1	467	2	D84938	H+-transporting AT
5	33	76.7	1042	2	A57534	mucin 5AC (clone L
6	33	76.7	1056	2	A54767	mucin MUC5B, trach
7	33	76.7	1384	2	A55184	tubulin-2 precurso
8	32	74.4	2229	2	T16199	hypothetical prote
9	31	72.1	279	2	AB1540	transcription requ
10	31	72.1	336	2	S47436	JpaN protein - Sal
11	31	72.1	336	2	AB0851	surface presentati
12	31	72.1	452	2	S22799	major capsid prote
13	31	72.1	478	2	E87563	histidinol-phospha
14	31	72.1	498	2	S13269	translation initia
15	31	72.1	406	1	F1MS4A	translation initia
16	31	72.1	406	2	S43681	translation initia
17	30	69.8	107	2	A72529	hypothetical prote
18	30	69.8	146	2	B90259	hypothetical prote
19	30	69.8	160	2	A82214	hypothetical prote
20	30	69.8	194	2	B90473	hypothetical prote
21	30	69.8	195	2	G90271	hypothetical prote
22	30	69.8	323	2	T47574	embryonic abundant
23	30	69.8	328	2	E86714	quinone oxidoreduc
24	30	69.8	370	1	A26998	probable phosphose
25	30	69.8	371	1	D76186	transcription fact
26	30	69.8	424	2	E88196	protein ZK1127.6 [
27	30	69.8	459	2	T29945	hypothetical prote
28	30	69.8	460	2	T29946	hypothetical prote
29	30	69.8	469	2	T40429	ferredoxin - NAD(+
30	30	69.8	472	2	AD1609	6-phospho-1-ketofr
31	30	69.8	472	2	AB1246	6-phosphofructokin
32	30	69.8	514	2	D71158	hypothetical protein
33	30	69.8	524	2	T08931	hypothetical prote
34	30	69.8	684	2	G90473	hypothetical prote
35	30	69.8	686	2	A38235	microtubule-associ
36	30	69.8	733	2	A45301	microtubule associ
37	30	69.8	918	2	B84580	hypothetical prote
38	30	69.8	946	2	F88196	protein ZK1127.6 [
39	30	69.8	1026	2	S35882	hypothetical prote
40	30	69.8	1414	2	T09481	mating type silenc
41	30	69.8	1630	2	S64403	ESP1 protein - yea
42	29	67.4	160	2	S47661	topoisomerase II -
43	29	67.4	245	2	T20234	hypothetical prote
44	29	67.4	445	2	E86202	hypothetical prote
45	29	67.4	459	2	S48368	probable membrane

ALIGNMENTS

RESULT 1
S51427
heparin-binding glycoprotein 88K - pig
C;Species: Sus scrofa domestica (domestic pig)
C;Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 22-Jun-1999
C;Accession: S51427
R;Shackelton, L.M.; Mann, D.M.; Millis, A.J.T.
submitted to the EMBL Data Library, January 1995
A;Description: Identification of a 88kDa heparin-binding glycopr[illegible] (ap-88K) in diff
A;Reference number: S51427
A;Accession: S51427
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-383 <SHA>
A;Cross-references: EMBL:Z47803; NID:g634097; PIDN:CAA87764.1; PID:g634098
C;Superfamily: Streptomyces chitinase chi40

Query Match 100.0%; Score 43; DB 2; Length 383;
Best Local Similarity 100.0%; Pred. No. 0.066;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FSKIASNIQ 9
||||||||| |
Db 106 FSKIASNIQ 114

RESULT 2
A49562
cartilage glycoprotein gp39 precursor - human
N;Alternate names: 39K synovial protein
C;Species: Homo sapiens (man)
C;Date: 23-Mar-1995 #sequence_revision 23-Mar-1995 #text_change 22-Jun-1999
C;Accession: A49562, S10677, A55162
R;Hakala, B.E.; White, C.; Recklies, A.D.
J. Biol. Chem. 268, 25803-25810, 1993
A;Title: Human cartilage gp-39, a major secretory product of articular chondrocytes a
A;Reference number: A49562; MUID:94064658
A;Accession: A49562
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-383 <HAK>
A;Cross-references: GB:M80927; NID:g348911; PIDN:AAA16074.1; PID:g348912
R;Nyirkos, P.; Golds, E.E.
Biochem. J. 269, 265-268, 1990
A;Title: Human synovial cells secrete a 39 kDa protein similar to a bovine mammary pr
A;Reference number: S10677; MUID:90328984
A;Accession: S10677
A;Molecule type: protein
A;Residues: 22-40,'X',42-45 <NYZ>
C;Superfamily: Streptomyces chitinase chi40
C;Keywords: cartilage; extracellular protein; glycoprotein

F;1-21/Domain: signal sequence #status predicted -SIG-
F;2-384/Product: cartilage glycoprotein gp39 #status predicted -MAT-

Query Match 100.0%; Score 43; DB 2; Length 384;
Best Local Similarity 100.0%; Pred. No. 0.066;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FSKIASNIQ 9
||||||||| |
Db 106 FSKIASNIG 114

RESULT 3
S61551
breast regressing protein brp39 precursor - mouse
C;Species: Mus musculus (house mouse)
C;Date: 19-Mar-1997 #sequence_revision 25-Apr-1997 #text_change 25-Jan-2002
C;Accession: S61551; S61550; I48271
R;Morrison, B.W.; Loder, F.
Oncogene 9, 3417-3426, 1994
A;Title: neu and ras initiate murine mammary tumors that share genetic markers generally
A;Reference number: I48271; MUID:95060797
A;Accession: S61551
A;Status: nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1-405 <MOR1>
A;Cross-references: EMBL:X94035; NID:g1085065; PIDN:CAA63031.1; PID:g1085066
R;Morrison, B.W.
submitted to the EMBL Data Library, November 1996
A;Reference number: S61550
A;Accession: S61550
A;Molecule type: mRNA
A;Residues: 1-245,'T',247-330,'H',332-459,'MVWALLCIFLSSIPSPRRTTPDTNNIRGALA' 488
A;Cross-references: EMBL:X94035; NID:g1085065; PIDN:CAA63031.1; PID:g1085066
A;Note: the differences at the carboxyl end are due to a frameshift error
C;Genetics:
A;Gene: brp39
C;Superfamily: Streptomyces chitinase ch140
F;1-21/Domain: signal sequence #status predicted -SIG-
F;22-405/Product: breast regressing protein brp39 #status predicted -MAT-

Query Match 83.7%; Score 36; DB 2; Length 405;
Best Local Similarity 77.8%; Pred. No. 2.9;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FSKIASNIQ 9
|| ||||||
Db 107 FSEIASNIE 115

RESULT 4
D84938
BC transporting ATP synthase (EC 3.6.1.34), flagellum-specific, imp protein - Buchnera sp.
C;Species: Buchnera sp.
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 23-Mar-2001
C;Accession: D84938
R;Shigenobu, S.; Watanabe, H.; Hattori, M.; Sakaki, Y.; Ishikawa, H.
Nature 407, 81-86, 2000
A;Title: Genome sequence of the endocellular bacterial symbiont of aphids Buchnera sp. A
A;Reference number: AB4930; MUID:20445173
A;Accession: D84938
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-467 <SHI>
A;Cross-references: GB:AP000398; GSPDB:GN00144
A;Experimental source: strain APS
C;Genetics:
A;Gene: flil; BU076
C;Superfamily: BC transporting ATP synthase alpha chain; BC transporting ATP synthase al
C;Keywords: hydrolase

Query Match 79.1%; Score 34; DB 2; Length 467;
Best Local Similarity 87.5%; Pred. No. 5.9;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FSKIASNI 8
|:||||:
Db 109 FSKIASNS 116

RESULT 5
A57584
mucin 5AC (clone 134, - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 08-Feb-1996 #sequence_revision 08-Feb-1996 #text_change 2-Apr-2
C;Accession: A57584
R;Lesuffleur, T.; Roche, F.; Hill, A.S.; Lacasa, M.; Porz, M.; Swal... M...
J. Biol. Chem. 270, 13665-13674, 1995
A;Title: Characterization of a mucin cDNA clone isolated from HT-29 ...
A;Reference number: A57584; MUID:95299957
A;Accession: A57584
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-1042 <LES>
A;Cross-references: GB:Z48434; NID:g1054767; PIDN:CAA88...; PID:g...
C;Genetics:
A;Gene: GDB:MUC5AC
A;Cross-references: GB:M64347; MIM:158377
A;Map position: 11p15.5-11p15.5
C;Superfamily: von Willebrand factor type C repeat homology
F;678-746/Domain: von Willebrand factor type C repeat homology <VW>

Query Match 74.2%; Score 33; DB 2; Length 1042;
Best Local Similarity 66.7%; Pred. No. 41;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FSKIASNIQ 9
| : |:||:
Db 448 FSKFANNIE 456

RESULT 6
A53767
mucin MUC5B, tracheobronchial - human
C;Species: Homo sapiens (man)
C;Date: 06-Jan-1995 #sequence_revision 08-Jan-1995 #text_change 2-Apr-2
C;Accession: A53767
R;Meerzaman, D.; Charles, P.; Daskal, E.; Polymeropoulos, M.H.; Mart... B.M.; Bero... M
J. Biol. Chem. 269, 12932-12939, 1994
A;Title: Cloning and analysis of cDNA encoding a major airway glycoprotein, human tra
A;Reference number: A53767; MUID:94240426
A;Accession: A53767
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-1056 <MEE>
A;Cross-references: GB:U07111; NID:g400072; PIDN:AAA...; PID:g4...
A;Note: authors translated the codon TAT for residue 1054 as Tha
C;Genetics:
A;Gene: GDB:MUC5B; MUC5
A;Cross-references: GDB:128436; OMIM:600770
A;Map position: 11p15.5-11p15.5
C;Superfamily: von Willebrand factor type C repeat homology
F;262-845/Domain: von Willebrand factor type C repeat homology <VW>

Query Match 74.2%; Score 33; DB 2; Length ...;
Best Local Similarity 66.7%; Pred. No. 41;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FSKIASNIQ 9
||| |:||:
Db 468 FSKFANNIE 476

RESULT 7
A55284
fibulin 2 precursor - human.
N;Alternate names: protein DKFZp586A1519.1
C;Species: Homo sapiens (man)
C;Date: 27-Jan-1995 #sequence_revision 27-Jan-1995 #text_change 31-Jul 2000
C;Accession: A55284; T 8744
R;Zhang, R.Z.; Pan, T.C.; Zhang, Z.Y.; Mattei, M.G.; Timpl, R.; Chu, M.L.
Genomics 22, 425-430, 1994
A;Title: Fibulin 2 (FBLN2). human cDNA sequence, mRNA expression, and mapping of the gen
A;Reference number: A55284; MUID:95104855
A;Accession: A55284
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-1184 <ZHE>
A;Cross-references: GB:X82494; NID:g535232; PIDN:CAA57876.1; PID:g535233
R;Wambutt, R.; Reinhardt, D.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, March 1999
A;Reference number: Z14471
A;Accession: I08744
A;Molecule type: mRNA
A;Residues: 656-719,'QLEPLMGAHDCSRKQFCVNITLGSFYCVNHTVLCAISYTLNAHRKCVD',720-853,'I',855-11
A;Cross-references: EMBL:AJ050095
A;Experimental source: adult uterus; clone DKFZp586A1519
C;Genetics:
A;Gene: GDB:FBLN2
A;Cross-references: GB:299037; OMIM:135821
A;Map position: 9q25-q34
A;Note: DKFZp586A1519.1
C;Superfamily: unassigned cEGF related proteins; EGF homology
C;Keywords: alternative splicing; extracellular matrix
F;1-27/Domain: signal sequence #status predicted <SIG>
F;28-1184/Product: fibulin 2 protein #status predicted <MAT>
F;905-941/Domain: EGF-like motif <EGF>

Query Match 76.7%; Score 23; DB 2; Length 1184;
Best Local Similarity 75.0%; Pred. No. 47;
Matches 6; Conservative 2; Mismatches 0; Indels 1; Gaps 0;

Qy 1 FSKIASNI 8
|:: | |
Db 905 FSQVASNT 912

RESULT 8
I16199
hypothetical protein F2B84.8 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
C;Accession: I16199
R;Leimbach, D.
submitted to the EMBL Data Library, December 1995
A;Description: The sequence of C.elegans cosmid F2B84.
A;Reference number: Z19475
A;Accession: I16199
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-2229 <LEI>
A;Cross-references: EMBL:U12844; NID:g1125755; PID:g1125756; PIDN:AAA84681.1; PESP:F2B84
C;Genetics:
A;Gene: CESP:F2B84.8
A;Introns: 44/3; 99/3; 147/3; 245/3; 340/3; 1321/3; 1471/3; 1723/3; 1777/3; 1845/3; 2

Query Match 71.4%; Score 22; DB 2; Length 2229;
Best Local Similarity 66.7%; Pred. No. 1.6e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FSKIASNIQ 9
:| |::: |
Db 1233 YSSIAANFQ 1241

RESULT 9
AH1540
transcription regulator homolog lin0864 [imported] - Listeria innocua (strain Clip112
C;Species: Listeria innocua
C;Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 27-Nov-2001
C;Accession: AH1540
R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloec
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi,
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A;Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.;
ok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehla
A;Title: Comparative genomics of Listeria species.
A;Reference number: AH1077; MUID:21537279; PMID:11679669
A;Accession: AH1540
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-279 <GLA>
A;Cross-references: GB:AL592022; PIDN:CAC96096.1; PID:g16413315; GSPDB:GN00178
A;Experimental source: strain Clip11262
C;Genetics:
A;Gene: lin0864

Query Match 72.1%; Score 31; DB 2; Length 279;
Best Local Similarity 66.7%; Pred. No. 28;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 FSKIASNIQ 9
| | : ||||
Db 80 FPKLVSNIQ 88

RESULT 10
S37306
spaN protein - Salmonella typhimurium
C;Species: Salmonella typhimurium
C;Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 08-Oct-1999
C;Accession: S37306; S60895
R;Groisman, E.A.; Ochman, H.
EMBO J. 12, 3779-3787, 1993
A;Title: Cognate gene clusters govern invasion of host epithelial cells by Salmonella
A;Reference number: S37304; MUID:94008985
A;Accession: S37306
A;Status: nucleic acid sequence not shown
A;Molecule type: DNA
A;Residues: 1-336 <GRO>
A;Cross-references: EMBL:X74525; NID:g404286; PIDN:CAA52193.1; PID:g404289
R;Collazo, C.M.; Zierler, M.K.; Galan, J.E.
Mol. Microbiol. 15, 25-38, 1995
A;Title: Functional analysis of the Salmonella typhimurium invasion genes invI and in
A;Reference number: S60893; MUID:95272391
A;Accession: S60895
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-336 <COL>
A;Cross-references: EMBL:U30472; NID:g507021; PIDN:AAA83441.1; PID:g507064
C;Genetics:
A;Gene: SpaN; InvJ

Query Match 72.1%; Score 31; DB 2; Length 336;
Best Local Similarity 77.8%; Pred. No. 34;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 FSKIASNIQ 9
|| | |:||
Db 286 FSLIPSNIQ 294

RESULT 11
AH0851
surface presentation of antigens protein (associated with type III secretion and virulen
C;Species: Salmonella enterica subsp. enterica serovar Typhi
A;Note: this species has also been called Salmonella typhi
C;Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 26-Nov-2001
C;Accession: AH0851
R;Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,
th, T.; Connerton, P.; Cronin, A.; Davies, P.; Parsley, N.M.; Doyd, L.; White, N.; Farrar,
, S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A;Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;
A;Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov
A;Reference number: AB0502; PMID:11677608
A;Accession: AH0851
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-336 <PAR>
A;Cross-references: GB:AL513382; PIDN:CAD05999.1; PID:g16503072; GSPDB:GN00170
C;Genetics:
A;Gene: spaN

Query Match 72.1%; Score 41; DB 2; Length 336;
Best Local Similarity 77.8%; Pred. No. 44;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 FSKIASNIQ 9
|||:|:|||
Db 286 FSLIFSNIQ 294

RESULT 12
S22799
major capsid protein gpN - phage P2
C;Species: phage P2
C;Date: 22-Nov-1994 #sequence_revision 10-Nov-1995 #text_change 09-Sep-1997
C;Accession: S22799; S16418
R;Linderoth, N.A.; Viermann, R.; Haggard-Ljungquist, E.; Christie, G.E.; Calendar, R.
Nucleic Acids Res. 19, 7207-7214, 1991
A;Title: Nucleotide sequence of the DNA packaging and capsid synthesis genes of bacterio
A;Reference number: S22796; MUID:92115571
A;Accession: S22799
A;Molecule type: DNA
A;Residues: 1-357 <LIN>
A;Cross-references: EMBL:X61229; NID:g15141; PID:g15145
C;Genetics:
A;Gene: N

Query Match 72.1%; Score 41; DB 2; Length 357;
Best Local Similarity 85.7%; Pred. No. 46;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FSKIASN 7
|||:|:|
Db 93 FSKLASN 99

RESULT 13
F87568
histidinol-phosphate aminotransferase [imported] - Caulobacter crescentus
C;Species: Caulobacter crescentus
C;Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 20-Apr-2001
C;Accession: F87568
R;Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.
B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A;Title: Complete Genome Sequence of Caulobacter crescentus.
A;Reference number: A87249; MUID:21173698; PMID:11259647
A;Accession: F87568
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-378 <SUN>
A;Cross-references: GB:AE006075; NID:g13419005; PIDN:AAK23113.1; GSPDB:GN00170
C;Genetics:
A;Gene: CC2544

Query Match 72.1%; Score 41; DB 2; Length 378;
Best Local Similarity 85.7%; Pred. No. 49;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FSKIASN 7
|||:|||
Db 50 FSKLASN 56

RESULT 14
S18269
translation initiation factor eIF-4A I - rabbit (fragment)
C;Species: Oryctolagus cuniculus (domestic rabbit)
C;Date: 19-Mar-1992 #sequence_revision 17-Apr-1998 #text_change 24-Feb-1999
C;Accession: S18269; S78482
R;Conroy, S.C.; Dever, T.E.; Owens, C.L.; Merrick, W.C.
Arch. Biochem. Biophys. 282, 364-371, 1990
A;Title: Characterization of the 46,000 dalton subunit of eIF-4F.
A;Reference number: S18269; MUID:90305155
A;Accession: S18269
A;Molecule type: mRNA
A;Residues: 5-394 <CON>
A;Experimental source: reticulocyte
A;Accession: S78482
A;Molecule type: protein
A;Residues: 1-25;38-53;col 102;142-170;188-194;219-225;235-266;273-347;col 349-<CON>
A;Experimental source: reticulocyte
C;Comment: This is one of the many factors involved in the complex processes of protein
a protein complex involved in "cap" recognition.
C;Genetics:
A;Gene: eIF-4A1
C;Function:
A;Description: translation initiation.
C;Superfamily: translation initiation factor eIF-4A
C;Keywords: ATP; blocked amino end; nucleotide binding; RNA-dependent ATPase; ribosome
F;1-398/Product: translation initiation factor eIF-4A.1 (fragment) #status experimental
F;68-75/Region: nucleotide-binding motif A (P-loop)
F;170-175/Region: nucleotide-binding motif B
F;174-177/Region: DEAD motif
F;74/Binding site: ATP (Lys) #status predicted

Query Match 72.1%; Score 41; DB 2; Length 398;
Best Local Similarity 66.7%; Pred. No. 51;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 FSKIASNIQ 9
|:|:||||
Db 192 FQKINSNIQ 200

RESULT 15
FIMS4A
translation initiation factor eIF-4A I - mouse
C;Species: Mus musculus (house mouse)
C;Date: 31-Mar-1988 #sequence_revision 29-Apr-1988 #text_change 16-Jan-
C;Accession: JS0039; A24267; B42267; S00005; A26267
R;Reddy, N.S.; Rath, W.W.; Braun, E.W.; Wahba, A.J.
Gene 70, 241-243, 1988
A;Title: Isolation and mapping of a gene for protein synthesis initiation factor 4A-I
A;Reference number: JS0039; MUID:89108007
A;Accession: JS0039
A;Molecule type: DNA
A;Residues: 1-299,'G',301-406 <RED>
A;Cross-references: GB:L36611; NID:g556406; PIDN:AAA50407.1; PID:g556
A;Note: the authors translated the codon GAG for residue 144 as Gly

A:Note: The codon given at bp 350 for (CH) is inconsistent with the authors' translation.
R:Nielsen, P.J.; McMaster, G.K.; Trachsel, H.
Nucleic Acids Res. 13, 6867, 1985
A:Title: Cloning of eukaryotic protein synthesis initiation factor genes: isolation and
A:Reference number: A42667; MUID:86041884
A:Accession: A24267
A:Molecule type: mRNA
A:Residues: 17-406 <N.E2>
A:Cross-references: GB:X03408; NID:g50813; PIDN:CAA27684.1; PID:g50813; GB:X03410; NID:g
A:Accession: B24267
A:Molecule type: mRNA
A:Residues: 37-406 <N.E2>
A:Cross-references: GB:X03410; NID:g50815; PIDN:CAA27684.1; PID:g50815; GB:X03408; NID:g
R:Nielsen, P.J.; Trachsel, H.
EMBO J. 7, 2097-2105, 1988
A:Title: The mouse protein synthesis initiation factor 4A gene family includes two relat
A:Reference number: S00698; MUID:88305098
A:Accession: S00698
A:Status: not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 'M',1-54;'V',197-406 <NIE3>
A:Cross-references: EMBL:X14421; NID:g50810
A:Accession: A59870
A:Status: not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-8 <NIE4>
A:Cross-references: EMBL:X14421; NID:g50810; PIDN:CAA32984.1; PID:g581664
C;Comment: This is one of the many factors involved in the complex assembly process of a
 a protein complex is called in "cap" recognition.
C;Genetics:
A;Gene: MGI:Eif4a1
A;Cross-references: MGI:95403
A;Map position: 11:49.0
A;Introns: 84-141 <142-194> 195-252 <253-305> 306-362 <363-405>
C;Superfamily: translation initiation factor eIF-4A
C;Keywords: ATP; nucleotide binding; P-loop; protein biosynthesis; RNA binding
F;76-83/Region: nucleotide-binding motif A (P-loop)
F;178-184/Region: nucleotide-binding motif B
F;182-185/Region: DEAD motif
F;82/Binding site: ATP (Lys) #status predicted

  Query Match 7.2 10; Score 81; DB 1; Length 406;
  Best Local Similarity 66.7%; Pred. No. 42;
  Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 FSKIASNIQ 9
  |:|:|||
Db 200 FQKLNSNIQ 208

Search completed: August 26, 2002, 4:15:48
Job time: 1.9 secs

GenCore version 4.5
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 26, 2002, 06:20:01 ; Search time 5.26 Seconds
(without alignments)
4.080 Million cell updates/sec

Title: US-09-744-282-4
Perfect score: 43
Sequence: 1 PSKIASNTL 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 39216560 residues

Total number of hits satisfying chosen parameters: 105221

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt 40:*

Prob. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	43	100.0	383	1	CH3L1_HUMAN	P36222 homo sapien
2	36	83.7	381	1	CH3L1_MOUSE	Q61362 mus musculu
3	34	79.1	467	1	PLL1_BOVAI	P57178 buchnera ap
4	33	76.7	1056	1	MUC5_HUMAN	P98088 homo sapien
5	33	76.7	1184	1	FIL_HUMAN	P21333 homo sapien
6	32	74.4	255	1	MOBC_BPT4	P04511 bacteriopha
7	31	72.1	338	1	SLAN_SALTY	P24076 salmonella
8	31	72.1	457	1	VIN_BPT2	P26777 bacteriopha
9	31	72.1	894	1	TF41_RABIT	P29662 oryctolagus
10	31	72.1	400	1	TF41_HUMAN	P04765 homo sapien
11	31	72.1	911	1	HVIE_PASMU	P57967 pasteurella
12	30	69.8	168	1	PIME_PLEOS	P81655 pleurotus o
13	30	69.8	870	1	SERC_RABIT	P10658 oryctolagus
14	30	69.8	732	1	IAU_MOUSE	P19837 mus musculu
15	30	69.8	751	1	IAU_RAT	P19832 rattus norv
16	30	69.8	2630	1	ESP1_YEAST	Q00684 saccharomyc
17	29	67.4	121	1	RK14_CYAPA	Q91142 cyanidium c
18	29	67.4	296	1	COPE_YEAST	P40509 saccharomyc
19	29	67.4	870	1	SERC_HUMAN	Q99727 homo sapien
20	29	67.4	885	1	IEBE_XYN	[illegible]
21	29	67.4	176	1	IAA4_LYCES	Q43868 lycopersico
22	29	67.4	196	1	IAA1_ARATH	Q03250 arabidopsis
23	29	67.4	207	1	KMSI_ECOLI	P11447 escherichia
24	29	67.4	428	2	YVPC_BACSU	P39611 bacillus
25	29	67.4	985	1	MDL1_CANAL	P87998 candida alb
26	29	67.4	809	1	SEVA_MYCGE	P47428 mycoplasma
27	29	67.4	809	1	SEVA_MYCPN	P75359 mycoplasma
28	29	67.4	1229	1	PMP1_CHLEN	Q92899 chlamydia p
29	29	67.4	1879	1	IBP1_YEAST	P41940 saccharomyc
30	29	67.4	2398	1	TOP2_PLAFK	P43641 plasmodium
31	29	67.4	2454	1	KIF3C_DROME	Q76123 drosophila
32	28	65.1	186	1	YG5R_YEAST	P53130 saccharomyc
33	28	65.1	230	1	RNH2_METJA	Q57599 methanococc
34	28	65.1	357	1	IGN3_RAT	P19814 rattus norv
35	28	65.1	370	1	SERC_MOUSE	Q99K85 mus musculu
36	28	65.1	451	1	S3K_MOUSE	Q99K70 mus musculu
37	28	65.1	453	1	CAKA_DRIVE	P87184 driveria [illegible]
38	28	65.1	568	1	BUN1_MOUSE	Q08539 mus musculu
39	28	65.1	588	1	BIN1_RAT	Q08839 rattus norv
40	28	65.1	593	1	BIN1_HUMAN	O00499 homo sapien
41	28	65.1	1401	1	LATA_LAIMA	P29641 laimodectus
42	28	65.1	2009	1	SEC7_YEAST	P11075 saccharomyc
43	28	65.1	3140	1	DPOZ_HUMAN	P56724 homo sapien
44	28	65.1	3418	1	BRC2_HUMAN	P51587 homo sapien
45	27	62.8	101	1	R14B_STRCO	Q98989 streptomyce

ALIGNMENTS

```
RESULT  1
CH3L1_HUMAN
ID   CH3L1_HUMAN     STANDARD;      PRT;   383 AA.
AC   P36222; P30923;
DT   01-JUL-1993 (Rel. 25, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Chitinase 3-like protein 1 precursor (Cartilage glycoprotein-39) (39-
DE   kDa) (39 kDa synovial protein) (YKL-40).
GN   CHI3L1.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC   TISSUE=Cartilage;
RX   MEDLINE 94064658; PubMed 8245017;
RA   Hakala B.E., White C., Recklies A.D.;
RT   "Human cartilage gp-39, a major secretory product of articular
RT   chondrocytes and synovial cells, is a mammalian member of a chitinase
RT   protein family.";
RL   J. Biol. Chem. 268:25803-25810(1993).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Blood;
RX   MEDLINE 97386591; PubMed 9244440;
RA   Kohli M., Krause S.W., Andreesen R.;
RT   "Molecular characterization of the gene for human cartilage gp-39
RT   (CHI3L1), a member of the chitinase protein family and marker of
RT   late stages of macrophage differentiation.";
RL   Genomics 43:221-225(1997).
RN   [3]
RP   SEQUENCE OF 22-45.
RX   MEDLINE 90328984; PubMed 2375755;
RA   Nyirkos P., Golds E.E.;
RT   "Human synovial cells secrete a 39 kDa protein similar to a bovine
RT   mammary protein expressed during the non-lactating period.";
RL   Biochem. J. 269:265-268(1990).
CC   -!- FUNCTION: MAY PLAY AN IMPORTANT ROLE IN THE CAPACITY OF CELLS TO
CC       RESPOND TO AND COPE WITH CHANGES IN THEIR ENVIRONMENT.
CC   -!- SUBCELLULAR LOCATION: Extracellular.
CC   -!- TISSUE SPECIFICITY: PRESENT IN ARTICULAR CHONDROCYTES, SYNOVIAL
CC       CELLS AS WELL AS IN LIVER. UNDETECTABLE IN MUSCLE, TISSUES OF
CC       PANCREAS, MONONUCLEAR CELLS, OR FIBROBLASTS.
CC   -!- PTM: GLYCOSYLATED.
CC   -!- SIMILARITY: BELONGS TO FAMILY 18 OF GLYCOSYL HYDROLASES.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
```

```
DR   EMBL; M86962; AAA36074.1; -.
DR   EMBL; Y08374; CAA69661.1; -.
DR   EMBL; Y08375; CAA69661.1; JOINED.
DR   EMBL; Y08376; CAA69661.1; JOINED.
DR   EMBL; Y08377; CAA69661.1; JOINED.
DR   EMBL; Y08378; CAA69661.1; JOINED.
DR   PIR; S10677; S10677.
DR   PIR; A33162; A33162.
DR   MIM; 601525; -.
DR   InterPro; IPR001579; Chitinase_2.
DR   InterPro; IPR001223; Glyco_hydro_18.
DR   Pfam; PF00704; Glyco_hydro_18; 1.
DR   PROSITE; PS01095; CHITINASE_18; FALSE_NEG.
KW   Glycoprotein; Signal.
FT   SIGNAL       1     21
FT   CHAIN       22    383       CHITINASE 3 LIKE PROTEIN 1.
FT   CARBOHYD    60     60       N-LINKED (GLCNAC...) (POTENTIAL).
SQ   SEQUENCE   383 AA;  42613 MW;  76ADDB298EEEC2D1 CRC64;


  Query Match             100.0%;  Score 43;  DB 1;  Length 383;
  Best Local Similarity   100.0%;  Pred. No. 0.031;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 FSKIASNIQ 9
           |||||||||
Db     106 FSKIASNIQ 114


RESULT   2
CHI1_MOUSE
ID   CHI1_MOUSE      STANDARD;      PRT:   381 AA.
AC   Q61362;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Chitinase 3 like protein 1 precursor (Cartilage glycoprotein-39)
DE   (GP-39) (BRP39 protein).
GN   CHI3L1 OR BRP39.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciuromorpha; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN FVB/N; TISSUE=Breast;
RX   MEDLINE=95360767; PubMed=7670700;
RA   Morrison B.W., Leder P.;
RT   "neu and ras initiate murine mammary tumors that share genetic markers
RT   generally absent in c-myc and int-2-initiated tumors.";
RL   Oncogene 9:3417-3426(1994).
CC   -!- FUNCTION: MAY PLAY AN IMPORTANT ROLE IN THE CATALYTIC ACTIVITIES
CC       RESPONSIBLE AND SPECIFIC WITH CHANGES IN THEIR ENVIRONMENT (BY
CC       SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: Extracellular (By similarity).
CC   -!- SIMILARITY: BELONGS TO FAMILY 18 OF GLYCOSYL HYDROLASES.
CC   -----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----------------------------------------------------------------------
DR   EMBL; X93035; CAA64803.1; -.
DR   MGD; MGI:1340899; Chi3l1.
DR   InterPro; IPR001579; Chitinase_2.
DR   InterPro; IPR001223; Glyco_hydro_18.
DR   Pfam; PF00704; Glyco_hydro_18; 1.
DR   PROSITE; PS01095; CHITINASE_18; FALSE_NEG.
KW   Glycoprotein; Signal.
FT   SIGNAL       1     21       POTENTIAL.
FT   CHAIN       22    381       CHITINASE 3 LIKE PROTEIN 1.
FT   CARBOHYD    60     60       N-LINKED (GLCNAC...) (POTENTIAL).
SQ   SEQUENCE   381 AA;  43003 MW;  EF6581EB1B4F045C CRC64;


  Query Match             88.7%;  Score 36;  DB 1;  Length 381;
  Best Local Similarity   77.8%;  Pred. No. 1.2;
  Matches    7;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 FSKIASNIQ 9
           |:|||:|||
Db     107 FSEIASNIE 115


RESULT   3
FLII_BUCAI
ID   FLII_BUCAI      STANDARD;      PRT:   467 AA.
AC   P57178;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
DE   Flagellum-specific ATP synthase (EC 3.6.3.14).
GN   FLII OR BU076.
OS   Buchnera aphidicola (subsp. Acyrthosiphon pisum) (Acyrthosiphon p...
OS   symbiotic bacterium).
OC   Bacteria; Proteobacteria; gamma subdivision; Buchnera.
OX   NCBI_TaxID=118099;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN TOKYO 1998;
RX   MEDLINE=20445173; PubMed=10993077;
RA   Shigenobu S., Watanabe H., Hattori M., Sakaki Y., Ishikawa H.;
RT   "Genome sequence of the endocellular bacterial symbiont of aphids
RT   Buchnera sp. APS.";
RL   Nature 407:81-86(2000).
CC   -!- FUNCTION: PROBABLE CATALYTIC SUBUNIT OF A PROTEIN TRANSLOCASE P-
CC       FLAGELLUM SPECIFIC EXPORT, OR A PROTON TRANSLOCASE INVOLVED IN
CC       LOCAL CIRCUITS AT THE FLAGELLUM, MAY BE INVOLVED IN A SPECIALIZED
CC       PROTEIN EXPORT PATHWAY THAT PROCEEDS WITHOUT SIGNAL PEPTIDE
CC       CLEAVAGE (BY SIMILARITY).
CC   -!- CATALYTIC ACTIVITY: ATP + H(2)O + H(+)(In) = ADP + phosphate +
CC       H(+)(Out).
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic (Potential).
CC   -!- SIMILARITY: TO OTHER TRANSPORT APPARATUS ATPASES AND F-1 ATPASE
CC       BETA SUBUNIT.
CC   -----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----------------------------------------------------------------------
DR   EMBL; AP001118; BAB12796.1; -.
DR   InterPro; IPR000194; ATPase_alpha_beta.
DR   Pfam; PF00006; ATP-synt_ab; 1.
DR   PROSITE; PS00152; ATPASE_ALPHA_BETA; 1.
KW   Hydrolase; Hydrogen ion transport; ATP synthesis; ATP binding;
KW   Transport; Protein transport; Flagella; Complete proteome.
FT   NP_BIND     195    202       ATP (BY SIMILARITY).
SQ   SEQUENCE   467 AA;  51948 MW;  0C87A34C1E29A7A2 CRC64;


  Query Match             79.1%;  Score 34;  DB 1;  Length 467;
  Best Local Similarity   87.5%;  Pred. No. 4.2;
  Matches    7;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 FSKIASNI 8
           ||||||||
Db     109 FSKIASNS 116
```

```
RESULT   4
MUC5_HUMAN
ID   MUC5_HUMAN     STANDARD;      PRT;  5074 AA.
AC   P98088;
DT   01-FEB-1996 (Rel. 33, Created)
DT   01-FEB-1996 (Rel. 33, Last sequence update)
DT   15-OCT-2001 (Rel. 40, Last annotation update)
DE   Tracheobronchial mucin (TBM) (Major airway glycoprotein) (Fragment).
GN   MUC5AC OR MUC5.
OS   Homo sapiens (Human)
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A., AND SEQUENCE OF 23-44, 67-76 AND 4??-506.
RC   TISSUE=Lung
RX   MEDLINE=94140378; PubMed=7511266;
RA   Meerzaman D., Charles P.J., Lasser E., Polymeropoulos M.H.,
RA   Martin B.M., Rose M.C.;
RT   "Cloning and analysis of cDNA encoding a major airway glycoprotein,
RT   human tracheobronchial mucin (MUC5).";
RL   J. Biol. Chem. 269:12832-12839(1994).
CC   -!- FUNCTION: MAJOR AIRWAY GLYCOPROTEIN.
CC   -!- SUBUNIT: MULTIMERIC.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- TISSUE SPECIFICITY: TRACHEOBRONCHIAL MUCOSAE (NOT GASTRIC).
CC   -!- SIMILARITY: TO VARIOUS MUCINS.
CC   -!- SIMILARITY: CONTAINS 1 VWFC DOMAIN.
CC   -!- SIMILARITY: CONTAINS 2 VWFD DOMAINS.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; U60711; AAA18481.1; -.
DR   MIM; 158374; -.
DR   InterPro; IPR000436; Sys_Rich.
DR   InterPro; IPR001007; VWFC.
DR   InterPro; IPR001846; VWD.
DR   Pfam; PF00007; Cys_Rich; 1.
DR   Pfam; PF00094; vwd; 2.
DR   SMART; SM00041; CT; 2.
DR   SMART; SM00214; VWC; 2.
DR   SMART; SM00216; VWD; 1.
DR   PROSITE; PS01185; CTCK_1; 1.
DR   PROSITE; PS01225; CTCK_2; 1.
DR   PROSITE; PS01208; VWFC; 1.
KW   Repeat; Glycoprotein
FT   NON_TER       1      1
FT   DOMAIN      149    521       1 X 1 AA TANDEM REPEATS OF T T T V T T/S.
FT   REPEAT      149    153       1.
FT   REPEAT      531    535       2.
FT   REPEAT      536    540       3.
FT   REPEAT      541    545       4.
FT   REPEAT      546    550       5.
FT   DOMAIN      661    730       VWFC.
FT   DOMAIN      917   1015       CTCK.
FT   DISULFID    917    967       BY SIMILARITY.
FT   DISULFID    942    981       BY SIMILARITY.
FT   DISULFID    946    997       BY SIMILARITY.
FT   DISULFID    964    999       BY SIMILARITY.
FT   DISULFID    966   1004       BY SIMILARITY.
FT   CARBOHYD    256    256       N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD    443    443       N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD    479    479       N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD    621    621       N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD    732    732       N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD    752    752       N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD    771    771       N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD    913    913       N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD    976    976       N-LINKED (GLCNAC...) (POTENTIAL).
FT   REGION      121      1       To A TandemRepeat Region in DOMAIN 1
FT                                HUMAN INTESTINAL MUCIN (MUC2).
FT   SIMILAR     113   1024       TO CYSTEINE-RICH DOMAIN OF
FT                                HUMAN INTESTINAL MUCIN (MUC2).
FT   SIMILAR     215    729       TO RAT MUCIN-LIKE PROTEIN.
FT   SIMILAR     223    705       TO BOVINE SUBMAXILLARY MUCIN.
FT   SIMILAR     777   1009       TO PORCINE SUBMAXILLARY MUCIN.
SQ   SEQUENCE   1056 AA;  113043 MW;  2F07CBA5C241FA02 CRC64;


  Query Match             76.7%;  Score 33;  DB 1;  Length 1056;
  Best Local Similarity   66.7%;  Pred. No. 16;
  Matches    6;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 FSKIASNIQ 9
           ||| |:|||
Db     438 FSKFANNIE 446


RESULT   5
FBL2_HUMAN
ID   FBL2_HUMAN     STANDARD;      PRT;  1184 AA.
AC   P98095;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Fibulin-2 precursor.
GN   FBLN2.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Fibroblast;
RX   MEDLINE=95104855; PubMed=7806230;
RA   Zhang R.Z., Pan T.C., Zhang Z.Y., Mattei M.G., Timpl R.,
RA   Chu M.-L.;
RT   "Fibulin-2 (FBLN2): human cDNA sequence, mRNA expression, and mapping
RT   of the gene on human and mouse chromosomes.";
RL   Genomics 22:425-430(1994).
CC   -!- FUNCTION: MAY BIND TO FIBRONECTIN AND SOME OTHER ECM
CC       CALCIUM DEPENDENT.
CC   -!- SUBUNIT: HOMOTRIMER; DISULFIDE LINKED.
CC   -!- SUBCELLULAR LOCATION: EXTRACELLULAR MATRIX. COMPONENT OF BOTH
CC       BASEMENT MEMBRANES AND OTHER CONNECTIVE TISSUES.
CC   -!- TISSUE SPECIFICITY: EXPRESSED IN HEART, PLACENTA AND OVARY.
CC   -!- SIMILARITY: CONTAINS 3 ANAPHYLATOXIN LIKE DOMAINS.
CC   -!- SIMILARITY: CONTAINS 11 EGF LIKE DOMAINS.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; X82494; CAA57876.1; -.
DR   HSSP; P07204; 1FGD.
DR   MIM; 135821; -.
DR   InterPro; IPR000020; Anaphylatoxin.
DR   InterPro; IPR000152; Asx_hydroxyl.
DR   InterPro; IPR000561; EGF-like.
DR   InterPro; IPR001881; EGF_Ca.
DR   Pfam; PF01821; ANATO; 2.
DR   Pfam; PF00008; EGF; 7.
DR   SMART; SM00104; ANATO; 3.
DR   SMART; SM00179; EGF_CA; 9.
DR   SMART; SM00181; EGF_like; 1.
DR   PROSITE; PS00010; ASX_HYDROXYL; 5.
```

```
DR   PROSITE; PS01186; ANAPHYLATOXIN 1; 3.
DR   PROSITE; PS01187; ANAPHYLATOXIN_2; 3.
DR   PROSITE; PS00022; EGF 1; FALSE_NEG.
DR   PROSITE; PS01186; EGF 2; 5.
DR   PROSITE; PS01187; EGF_CA; 9.
KW   Signal; Glycoprotein; Extracellular matrix; Plasma; EGF-like domain;
KW   Calcium binding; Repeat.
FT   SIGNAL        1     27       POTENTIAL.
FT   CHAIN        28   1184       FIBULIN-2.
FT   DOMAIN       28    444       N.
FT   DOMAIN       28    177       SUBDOMAIN NA (CYS-RICH).
FT   DOMAIN      178    444       SUBDOMAIN NB (CYS-FREE).
FT   DOMAIN      445    480       ANAPHYLATOXIN-LIKE 1.
FT   DOMAIN      488    519       ANAPHYLATOXIN-LIKE 2.
FT   DOMAIN      521    553       ANAPHYLATOXIN LIKE 3.
FT   DOMAIN      604    645       EGF-LIKE 1, CALCIUM BINDING.
FT   DOMAIN      679    718       EGF LIKE 2.
FT   DOMAIN      719    763       EGF-LIKE 3, CALCIUM BINDING.
FT   DOMAIN      764    809       EGF-LIKE 4, CALCIUM-BINDING.
FT   DOMAIN      816    857       EGF LIKE 5, CALCIUM BINDING.
FT   DOMAIN      858    900       EGF LIKE 6, CALCIUM BINDING.
FT   DOMAIN      901    942       EGF LIKE 7, CALCIUM BINDING.
FT   DOMAIN      944    981       EGF-LIKE 8, CALCIUM-BINDING.
FT   DOMAIN      982   1024       EGF-LIKE 9, CALCIUM BINDING.
FT   DOMAIN     1025   1069       EGF LIKE 10, CALCIUM BINDING.
FT   DOMAIN     1070   1184       DOMAIN III.
FT   DISULFID    445    472       BY SIMILARITY.
FT   DISULFID    446    479       BY SIMILARITY.
FT   DISULFID    459    480       BY SIMILARITY.
FT   DISULFID    488    518       BY SIMILARITY.
FT   DISULFID    502    519       BY SIMILARITY.
FT   DISULFID    521    545       BY SIMILARITY.
FT   DISULFID    522    552       BY SIMILARITY.
FT   DISULFID    535    553       BY SIMILARITY.
FT   DISULFID    608    620       BY SIMILARITY.
FT   DISULFID    616    629       BY SIMILARITY.
FT   DISULFID    631    644       BY SIMILARITY.
FT   DISULFID    683    693       BY SIMILARITY.
FT   DISULFID    689    702       BY SIMILARITY.
FT   DISULFID    704    717       BY SIMILARITY.
FT   DISULFID    723    736       BY SIMILARITY.
FT   DISULFID    730    745       BY SIMILARITY.
FT   DISULFID    751    762       BY SIMILARITY.
FT   DISULFID    768    781       BY SIMILARITY.
FT   DISULFID    775    790       BY SIMILARITY.
FT   DISULFID    796    808       BY SIMILARITY.
FT   DISULFID    814    827       BY SIMILARITY.
FT   DISULFID    827    836       BY SIMILARITY.
FT   DISULFID    843    856       BY SIMILARITY.
FT   DISULFID    862    875       BY SIMILARITY.
FT   DISULFID    869    884       BY SIMILARITY.
FT   DISULFID    886    899       BY SIMILARITY.
FT   DISULFID    905    917       BY SIMILARITY.
FT   DISULFID    913    926       BY SIMILARITY.
FT   DISULFID    928    941       BY SIMILARITY.
FT   DISULFID    947    956       BY SIMILARITY.
FT   DISULFID    952    965       BY SIMILARITY.
FT   DISULFID    967    980       BY SIMILARITY.
FT   DISULFID    986    998       BY SIMILARITY.
FT   DISULFID    994   1007       BY SIMILARITY.
FT   DISULFID   1009   1023       BY SIMILARITY.
FT   DISULFID   1029   1042       BY SIMILARITY.
FT   DISULFID   1036   1051       BY SIMILARITY.
FT   DISULFID   1056   1068       BY SIMILARITY.
FT   CARBOHYD    180    180       N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD    507    507       N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD   1035   1035       N-LINKED (GLCNAC...) (POTENTIAL).
SQ   SEQUENCE   1184 AA;  126543 MW;  CA48490A55F9E050 CRC64;


  Query Match             25.7%;  Score 43;  DB 1;  Length 1184;
  Best Local Similarity   75.0%;  Pred. No. 18;
  Matches    3;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 FSKIASNI 8
             ||:|:|:|
Db     665 FSQVASNI 672


RESULT   6
MOBD_BPT4
ID   MOBD_BPT4       STANDARD;      PRT:   259 AA.
AC   P39241; Q96216;
DT   01-FEB-1995 (Rel. 31, Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
DE   Probable mobile endonuclease 5.
GN   MOBD OR TK.-10.
OS   Bacteriophage T4.
OC   Viruses; dsDNA viruses, no RNA stage; Caudovirales; Myoviridae;
OC   T4-like phages.
OX   NCBI_TaxID=10665;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Mizhavia N., Marusich E., Sjarakhidze I., Neitzel J., Interes ...
RA   Awaya M., Eidemmi ter J., Calanda L., Tracy I., Gailbreath K.
RA   Paddison P., Anderson R., Stidham I., Blattner F., Kutter E.M.;
RT   "The 10.7 kb 'nonessential' region of bacteriophage T4 between tk
RT   genes tk and nrdC  twenty new T4 genes, generally conserved amon
RT   T-even phages.";
RL   Submitted (NOV 1996) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Kutter E., Arisaka F., Kanaiwa I., Isoguchi A., Mosig G.,
RA   Mesyanzhinov V., Kuger W., Stidham I., Thomas F.;
RT   "Bacteriophage T4 genome analysis.";
RL   Submitted (JUL 2000) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: TO PHAGE T4 MOBB AND MOBC.
CC   ---------- ------ ------ ----- ---- ---- ---- ---- ---- ---- ---- ---
CC   This SWISS-PROT entry is copyright. It is produced thr ugh a ... at ta' n
CC   between the Swiss Institute of Bioinformatics and the EMBL ... tat ...
CC   the European Bioinformatics Institute. There are no rest...ti... ... ...
CC   use by non-profit institutions as long as its content ... ... is ... way
CC   modified and this statement is not removed. Usage by and ... ... ...rcial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------- ------ ------ ----- ---- ---- ---- ---- ---- ---- ---- ---
DR   EMBL; U76612; AAB20970.1;  -.
DR   EMBL; AF158101; AAD42542.1;  -.
DR   InterPro; IPR003271; HNH.
DR   InterPro; IPR003615; HNH_nuc.
DR   Pfam; PF01844; HNH; 1.
DR   SMART; SM00507; HNHc; 1.
KW   Hydrolase; Nuclease; Endonuclease.
SQ   SEQUENCE   259 AA;  30452 MW;  C149B7C72EA108F4 CRC64;


  Query Match             24.4%;  Score 41;  DB 1;  Length 259;
  Best Local Similarity   87.5%;  Pred. No. 4.5;
  Matches    7;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 FSKIASNT 8
             |||| |||
Db     105 FSKIHSNT 112


RESULT   7
SPAN_SALTY
ID   SPAN_SALTY     STANDARD;      PRT:   595 AA.
AC   P40613;
DT   01-FEB-1995 (Rel. 31, Created)
DT   01-FEB-1995 (Rel. 31, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
DE   Surface presentation of antigens protein spaN (Invasion protein ...).
GN   SPAN OR INVJ OR SIM2892.
OS   Salmonella typhimurium.
```

```
OS   Bacteria; Proteobacteria; Gamma subdivision; Enterobacteriaceae;
OC   Salmonella.
OX   NCBI_TaxID=602;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=94209898; PubMed=8143844;
RA   Groisman E.A., Ochman H.;
RT   "Cognate gene clusters govern invasion of host epithelial cells by
RT   Salmonella typhimurium and Shigella flexneri.";
RL   EMBO J. 12:4779-4787(1993).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=SR-11;
RX   MEDLINE=95229...; PubMed=7775804;
RA   Collazo C., Zierler M.K., Galan J.E.;
RT   "Functional analysis of the Salmonella typhimurium invasion genes
RT   invI and invJ and identification of a target of the protein secretion
RT   apparatus encoded in the inv locus.";
RL   Mol. Microbiol. 15:25-38(1995).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 14028 / SGSC 2262;
RX   MEDLINE=20...; PubMed=11...;
RA   Boyd E.F., Li J., Nelson K., Ochman H., Selander R.K.;
RL   Submitted (JAN-2000) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   SEQUENCE FROM N.A.
RC   STRAIN=LT2 / SGSC1412 / ATCC 700720;
RX   MEDLINE=20...; PubMed=11...;
RA   McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
RA   Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
RA   Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
RA   Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
RA   Waterston R., Wilson R.K.;
RT   "Complete genome sequence of Salmonella enterica serovar Typhimurium
RT   LT2.";
RL   Nature 413:852-856(2001).
CC   -!- FUNCTION: INVOLVED IN A SECRETORY PATHWAY RESPONSIBLE FOR THE
CC       SURFACE PRESENTATION OF DETERMINANTS NEEDED FOR THE ENTRY OF
CC       SALMONELLA SPECIES INTO MAMMALIAN CELLS.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic (potential).
CC   -!- SIMILARITY: BELONGS TO THE SPAN FAMILY.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; X76...; CAA53...; -.
DR   EMBL; U1...; AAA8...; -.
DR   EMBL; M4...; AAC43...; -.
DR   EMBL; AE00...; AAL21...; -.
DR   PIR; S3...; S...
DR   SWISS-2DPAGE; Q...
DR   InterPro; IPR...; invas_inv.
DR   Pfam; PF...; SPAN; 1.
DR   PRINTS; PR...; SPANFAMILY.
KW   Virulence; Complete proteome.
SQ   SEQUENCE   ... AA;  ... MW;  ... CRC64;


  Query Match             ...%;  Score 31;  DB 1;  Length ...;
  Best Local Similarity   ...%;  Pred. No. 14;
  Matches    ...;  Conservative    ...;  Mismatches    ...;  Indels    0;  Gaps    0;

Qy       1 FSKIASNTQ 9
             || : |||
Db     286 FSLIPSNTQ 294


RESULT   8
VPN_BPP2
ID   VPN_BPP2       STANDARD;      PRT;   357 AA.
AC   P25477;
DT   01-MAY-1992 (Rel. 22, Created)
DT   01-MAY-1992 (Rel. 22, Last sequence update)
DT   01-OCT-1996 (Rel. 34, Last annotation update)
DE   Major capsid protein precursor (GPN).
GN   N.
OS   Bacteriophage P2.
OC   Viruses; dsDNA viruses, no RNA stage; Caudovirales; Myoviridae;
OC   P2-like viruses.
OX   NCBI_TaxID=10679;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=92115571; PubMed=1837455;
RA   Linderoth N.A., Ziermann R., Bagqaard-Ljunqquist E., Christie G.E.,
RA   Calendar R.;
RT   "Nucleotide sequence of the DNA packaging and capsid synthesis genes
RT   of bacteriophage P2.";
RL   Nucleic Acids Res. 19:7207-7214(1991).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91220668; PubMed=1840708;
RA   Fox E.W., Sandino M.G., Williams J., Bagqaard-Ljunqquist E.,
RA   Lindqvist B.H.;
RT   "Morphopoietic switch mutations of bacteriophage P2.";
RL   Virology 182:41-46(1991).
CC   -!- FUNCTION: P2 PROHEADS ARE CAPSIDS CONSIST PRIMARILY OF N*, A 9.7
CC       KDa PROTEIN AND TWO MINOR COMPONENTS, B1 (29 KDa) AND B2 (30.4
CC       KDa).
CC   -!- PTM: N*, B1 AND B2 ALL DERIVE FROM FULL LENGTH N PROTEIN BY AMINO-
CC       TERMINAL PROCESSING.
CC   -!- SIMILARITY: STRONG, TO PHAGE HP1 PROTEIN ORF18.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; AF063097; AAD03271.1; -.
DR   PIR; S16413; S16413.
DR   PIR; S22799; S22799.
KW   Coat protein.
SQ   SEQUENCE   357 AA;  40246 MW;  A203A0560A7EDB8B CRC64;


  Query Match             72.1%;  Score 31;  DB 1;  Length 357;
  Best Local Similarity   85.7%;  Pred. No. 15;
  Matches    6;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 FSKIASN 7
             ||:|||
Db     93 FSKLASN 99


RESULT   9
IF4I_RABIT
ID   IF4I_RABIT      STANDARD;      PRT;   394 AA.
AC   P29562;
DT   01-APR-1993 (Rel. 25, Created)
DT   01-APR-1993 (Rel. 25, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Eukaryotic initiation factor 4A I (eIF-4A I) (eIF4A I) (Fragment).
GN   EIF4A1 OR EIF4A.
OS   Oryctolagus cuniculus (Rabbit).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Laqomorpha; Leporidae; Oryctolaqus.
OX   NCBI_TaxID=9986;
RN   [1]
RP   SEQUENCE.
```

```
RX   MEDLINE 91058155; PubMed 2241157;
RA   Ozes O.N., Petryniak B., Kwon C.I., Merrick W.C.;
RT   "Characterization of the 46,000-dalton subunit of eIF 4F.";
RL   Arch. Biochem. Biophys. 282:363-371(1990).
CC   -!- FUNCTION: EIF-4A IS BOTH A SUBUNIT OF A HIGH MOLECULAR WEIGHT
CC       PROTEIN COMPLEX INVOLVED IN CAP RECOGNITION AND IS RELEASED AS A
CC       SINGLE POLYPEPTIDE CHAIN FOR MRNA BINDING (~ 410 AMINO ACIDS, 47 KD). IT IS AN
CC       ATP DEPENDENT SINGLE STRANDED DNA BINDING PROTEIN WITH A SEQUENCE
CC       INDEPENDENT UNWINDING ACTIVITY (HELICASE).
CC   -!- SIMILARITY: BELONGS TO THE DEAD BOX HELICASE FAMILY.
DR   PIR; S13266; S13269.
DR   HSSP; Q58083; 1HV8.
DR   InterPro; IPR000464; DEAD.
DR   InterPro; IPR000629; DEAD_ATP_helicase.
DR   InterPro; IPR001650; Helicase_C.
DR   Pfam; PF00270; DEAD; 1.
DR   Pfam; PF00271; helicase_C; 1.
DR   SMART; SM00487; DEXDc; 1.
DR   SMART; SM00490; HELICc; 1.
DR   PROSITE; PS00039; DEAD_ATP_HELICASE; 1.
KW   Initiation factor; Protein biosynthesis; ATP-binding; RNA-binding;
KW   DNA-binding; Helicase.
FT   NON_TER       1      1
FT   NP_BIND      64     71       ATP (BY SIMILARITY).
FT   SITE        171    173       DEAD-BOX.
SQ   SEQUENCE   394 AA;  44611 MW;  C07FD544C0B20722 CRC64;


  Query Match             72.1%;  Score 31;  DB 1;  Length 394;
  Best Local Similarity   66.7%;  Pred. No. 17;
  Matches    6;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 FSKIASNTQ 9
           | : ||| |
Db     188 FQKLNSNTQ 196


RESULT  10
IF4A_HUMAN
ID   IF4A_HUMAN     STANDARD;      PRT;   406 AA.
AC   P04765; Q13514;
DT   13-AUG-1987 (Rel. 05, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Eukaryotic initiation factor 4A-I (eIF-4A-I) (eIF4A-I).
GN   EIF1AI OR EIF4A OR DDX2A.
OS   Homo sapiens (Human), and
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606; 10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   SPECIES=Human;
RX   MEDLINE=86261841; PubMed=3493113;
RA   Kim N.S., Kato I., Abo N., Kato S.;
RT   "Nucleotide sequence of human cDNA encoding eukaryotic initiation
RT   factor 4AI.";
RL   Nucleic Acids Res. 21:2012-2012(1993).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   SPECIES=Mouse;
RX   MEDLINE=86041891; PubMed=3940589;
RA   Nielsen P.J., McMaster G.K., Trachsel H.;
RT   "Cloning of eukaryotic protein synthesis initiation factor genes:
RT   isolation and characterization of cDNA clones encoding factor
RT   eIF-4A.";
RL   Nucleic Acids Res. 13:6867-6880(1985).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   SPECIES=Mouse;
RX   MEDLINE=89198007; PubMed=3215517;
RA   Reddy N.S., Roth W.W., Bragg P.W., Wahba A.J.;
RT   "Isolation and mapping of a gene for protein synthesis initiation
RT   factor 4A and its expression during differentiation of Friend
RT   erythroleukemia cells.";
RL   Gene 70:231-243(1988).
RN   [4]
RP   SEQUENCE FROM N.A.
RC   SPECIES=Mouse; STRAIN=129;
RA   Miyashita A., Shimizu N., Nakajima T., Ishii S., Kimura A.;
RL   Submitted (MAR-1996) to the EMBL/GenBank/DDBJ databases.
RN   [5]
RP   MUTAGENESIS.
RC   SPECIES=Mouse;
RX   MEDLINE=92331617; PubMed=1378097;
RA   Pause A., Sonenberg N.;
RT   "Mutational analysis of a DEAD box RNA helicase: the mammalian
RT   translation initiation factor eIF-4A.";
RL   EMBO J. 11:2643-2654(1992).
CC   -!- FUNCTION: EIF-4A IS BOTH A SUBUNIT OF A HIGH MOLECULAR WEIGHT
CC       PROTEIN COMPLEX INVOLVED IN CAP RECOGNITION AND IS RELEASED AS A
CC       SINGLE POLYPEPTIDE CHAIN FOR MRNA BINDING (~ 410 AMINO ACIDS, 47 KD). IT IS AN
CC       ATP DEPENDENT SINGLE STRANDED DNA BINDING PROTEIN WITH A SEQUENCE
CC       INDEPENDENT UNWINDING ACTIVITY (HELICASE).
CC   -!- SIMILARITY: BELONGS TO THE DEAD BOX HELICASE FAMILY.
CC   -----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----------------------------------------------------------------------
DR   EMBL; J03248; BAA02897.1; -.
DR   EMBL; X03039; CAA26842.1; ALT_INIT.
DR   EMBL; X03039; CAA26843.1; ALT_INIT.
DR   EMBL; X03040; CAA26845.1; ALT_INIT.
DR   EMBL; X03040; CAA26846.1; ALT_INIT.
DR   EMBL; L36631; AAA50407.1; -.
DR   EMBL; M22873; AAA50407.1; JOINED.
DR   EMBL; L36608; AAA50407.1; JOINED.
DR   EMBL; L36609; AAA50407.1; JOINED.
DR   EMBL; L36610; AAA50407.1; JOINED.
DR   EMBL; AB011550; BAA25075.1; -.
DR   PIR; A24267; EIMS4A.
DR   PIR; S43681; S43681.
DR   PIR; S00986; S00986.
DR   HSSP; Q58083; 1HV8.
DR   MIM; 602641; -.
DR   MGD; MGI:95830; Eif4a1.
DR   InterPro; IPR001410; DEAD.
DR   InterPro; IPR000629; DEAD_ATP_helicase.
DR   InterPro; IPR001650; Helicase_C.
DR   Pfam; PF00270; DEAD; 1.
DR   Pfam; PF00271; helicase_C; 1.
DR   SMART; SM00487; DEXDc; 1.
DR   SMART; SM00490; HELICc; 1.
DR   PROSITE; PS00039; DEAD_ATP_HELICASE; 1.
KW   Initiation factor; Protein biosynthesis; ATP-binding; RNA-binding;
KW   DNA-binding; Helicase.
FT   NP_BIND      76     83       ATP (BY SIMILARITY).
FT   SITE        182    185       DEAD-BOX.
FT   CONFLICT    300    300       S -> C (IN REF. 5).
SQ   SEQUENCE   406 AA;  45154 MW;  6EFB0996F8041420 CRC64;


  Query Match             72.1%;  Score 31;  DB 1;  Length 406;
  Best Local Similarity   66.7%;  Pred. No. 17;
  Matches    6;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 FSKIASNTQ 9
           | : ||| |
Db     200 FQKLNSNTQ 208
```

```
RESULT  11
ILVD_PASMU
ID   ILVD_PASMU     STANDARD;      PRT:   611 AA.
AC   P57967;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Dihydroxy-acid dehydratase (EC 4.2.1.9) (DAD).
GN   ILVI OR PM1425.
OS   Pasteurella multocida.
OC   Bacteria; Proteobacteria; Gamma subdivision; Pasteurellaceae;
OC   Pasteurella.
OX   NCBI_TaxID=747;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=PM70;
RX   MEDLINE=21145866; PubMed=11248100;
RA   May B.J., Zhang Q., Li L.L., Paustian M.L., Whittam T.S., Kapur V.;
RT   "Complete genomic sequence of Pasteurella multocida Pm70.";
RL   Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).
CC   -!- CATALYTIC ACTIVITY: 2,3-dihydroxy-3-methylbutanoate = 3-methyl-2-
CC       oxobutanoate + H(2)O.
CC   -!- COFACTOR: BINDS 1 4FE-4S CLUSTER (POTENTIAL).
CC   -!- PATHWAY: FOURTH STEP IN VALINE AND ISOLEUCINE BIOSYNTHESIS.
CC   -!- SIMILARITY: BELONGS TO THE ILVD / EDD FAMILY.
CC   ---------- ---------- ---------- ---------- ----------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------- ---------- ---------- ---------- ----------
DR   EMBL; AE006100; AAK03709.1; -.
DR   InterPro; IPR000581; ILVD_EDD.
DR   Pfam; PF00920; ILVD_EDD; 1.
DR   Prodom; PD002691; ILVD_EDD; 1.
DR   PROSITE; PS00886; ILVD_EDD_1; 1.
DR   PROSITE; PS00887; ILVD_EDD_2; 1.
KW   Branched chain amino acid biosynthesis; Lyase; Iron-sulfur; 4Fe-4S;
KW   Complete proteome.
FT   METAL       12     12       IRON-SULFUR (4FE-4S) (POTENTIAL).
FT   METAL       105    105      IRON-SULFUR (4FE-4S) (POTENTIAL).
SQ   SEQUENCE   611 AA;  64472 MW;  6A3D804A3660761 CRC64;


  Query Match             21.1%;  Score 41;  DB 1;  Length 611;
  Best Local Similarity   75.0%;  Pred. No. 20;
  Matches    6;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       2 SKIASNIQV 9
           : | |||
Db     431 SKVAPNIQV 439


RESULT  12
PGMP_PLEOS
ID   PGMP_PLEOS     STANDARD;      PRT:   196 AA.
AC   P80656;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   15-JUL-1999 (Rel. 38, Last annotation update)
DE   Peptidyl-LYS metalloendopeptidase (EC 3.4.24.20) (MEP) (PGMEP).
OS   Pleurotus ostreatus (Oyster mushroom) (White-rot fungus).
OC   Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Basidiomycetidae;
OC   Agaricales; Pleurotaceae; Pleurotus.
OX   NCBI_TaxID=5322;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=98436068; PubMed=9744736;
RA   Nonaka T., Ishimaru N., Hashimoto Y., Takio K.;
RT   "Amino acid sequences of metalloendopeptidases specific for
RT   acyl-lysine bonds from Grifola frondosa and Pleurotus ostreatus
RT   fruiting bodies.";
RL   J. Biol. Chem. 272:30052-30059(1997).
CC   -!- CATALYTIC ACTIVITY: Preferential cleavage in proteins: Xaa-|-Lys
CC       (in which Xaa may be Pro).
CC   -!- COFACTOR: BINDS 1 ZINC ION.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- MASS SPECTROMETRY: MW=17927; METHOD=MALDI.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M35 (ZINC
CC       METALLOPROTEASE). ALSO KNOWN AS THE DEUTEROLYSIN SUBFAMILY.
DR   InterPro; IPR000130; Zn_MPpeptdse.
DR   PROSITE; PS00142; ZINC_PROTEASE; FALSE_NEG.
KW   Hydrolase; Metalloprotease; Zinc.
FT   DISULFID      6      76       BY SIMILARITY.
FT   DISULFID     78      98       BY SIMILARITY.
FT   METAL       116     116       ZINC (CATALYTIC) (BY SIMILARITY).
FT   ACT_SITE    119     119       BY SIMILARITY.
FT   METAL       122     122       ZINC (CATALYTIC) (BY SIMILARITY).
SQ   SEQUENCE   168 AA;  17925 MW;  78E5556CB1566547 CRC64;


  Query Match             69.8%;  Score 30;  DB 1;  Length 168;
  Best Local Similarity   75.0%;  Pred. No. 12;
  Matches    6;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 FSKIASNI 8
           || |:|:|
Db      61 FSSISSNT 68


RESULT  13
SERC_RABIT
ID   SERC_RABIT     STANDARD;      PRT:   370 AA.
AC   P10658;
DT   01-JUL-1989 (Rel. 11, Created)
DT   01-JUL-1989 (Rel. 11, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   Phosphoserine aminotransferase (EC 2.6.1.52) (PSAT) (Endometrial
DE   progesterone-induced protein) (EPIP).
GN   PSA.
OS   Oryctolagus cuniculus (Rabbit).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX   NCBI_TaxID=9986;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=88000617; PubMed=3651428;
RA   Mishrahi M., Atger M., Milgrom E.;
RT   "A novel progesterone-induced messenger RNA in rabbit and human
RT   endometria. Cloning and sequence analysis of the complementary DNA.";
RL   Biochemistry 26:3975-3982(1987).
RN   [2]
RP   POSSIBLE FUNCTION.
RX   MEDLINE=90045971; PubMed=2082527;
RA   van der Zel A., Lam H. M., Winkler M.E.;
RT   "Extensive homology between the Escherichia coli pdxF 5'-(PdxF)
RT   aminotransferase and a protein encoded by a progesterone induced mRNA
RT   in rabbit and human endometria.";
RL   Nucleic Acids Res. 17:8677-8679(1989).
CC   -!- CATALYTIC ACTIVITY: O-phospho-L-serine + 2-oxoglutarate =
CC       phosphonooxypyruvate + L-glutamate.
CC   -!- COFACTOR: PYRIDOXAL PHOSPHATE.
CC   -!- PATHWAY: REQUIRED BOTH IN MAJOR PHOSPHORYLATED PATHWAY OF SERINE
CC       BIOSYNTHESIS AND IN THE BIOSYNTHESIS OF PYRIDOXINE.
CC   -!- SUBUNIT: HOMODIMER.
CC   -!- INDUCTION: BY PROGESTERONE.
CC   -!- SIMILARITY: BELONGS TO CLASS V OF PYRIDOXAL PHOSPHATE DEPENDENT
CC       AMINOTRANSFERASES.
CC   ---------- ---------- ---------- ---------- ----------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
```

```
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed.  Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/
CC   or send an email to license@isb-sib.ch).
CC   ---------- ---------- ---------- ---------- ---------- ----------
DE   EMBL; M17232; AAA51215.1; -.
DR   PIR; A26998; A26998.
DR   HSSP; P23721; 1BIN.
DR   InterPro; IPR001152; Aminotransf_class_V.
DR   InterPro; IPR004838; Thosphser_aminotransf.
DR   Pfam; PF00266; aminotran_5; 1.
DR   ProDom; PD001544; Phosphser_aminotransf; 1.
DR   PROSITE; PS00595; AA_TRANSFER_CLASS_5; 1.
KW   Serine biosynthesis; Transferase; Aminotransferase;
KW   Pyridoxal phosphate.
FT   BINDING     203    230       PYRIDOXAL PHOSPHATE (BY SIMILARITY).
SQ   SEQUENCE   370 AA;  40621 MW;  7AB3A6E6B2D4085F CRC64;


  Query Match             69.8%;  Score 40;  DB 1;  Length 370;
  Best Local Similarity   55.6%;  Pred. No. 27;
  Matches    5;  Conservative    3;  Mismatches    1;  Indels    3;  Gaps    1;

Qy       1 ESKIASNIQ 9
           |||: :|||:
Db     49 EAKIVNNIE 57


RESULT  14
TAU_MOUSE
ID   TAU_MOUSE      STANDARD;      PRT;   732 AA.
AC   P10637; P10638; Q60084; Q60085; Q60086; Q62286;
DT   01-JUL 1989 (Rel. 11, Created)
DT   16-OCT 2001 (Rel. 40, Last sequence update)
DT   01-MAR 2002 (Rel. 41, Last annotation update)
DE   Microtubule-associated protein tau (Neurofibrillary tangle protein
DE   (Paired helical filament tau) (PHF-tau).
GN   MAPT OR MTAPT OR TAU.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.  (ISOFORM PNS TAU).
RC   TISSUE=Neuroblastoma;
RX   MEDLINE=92262443; PubMed=1374898;
RA   Couchie D., Mavilia C., Georgieff I.S., Liem R.K.H., Shelanski M.L.,
RA   Nunez J.;
RT   "Primary structure of high molecular weight tau present in the
RT   peripheral nervous system.";
RT   Proc. Natl. Acad. Sci. U.S.A. 89:4378-4381(1992).
RN   [2]
RP   SEQUENCE FROM N.A. (ISOFORMS TAU A, TAU C AND TAU E).
RC   STRAIN=Bim/SPF; TISSUE=Liver, Kidney, and Brain;
RX   MEDLINE=95012085; PubMed=7927211;
RA   Kenner L., el-Shabrawi Y., Hutter H., Forstner M., Pallua A.,
RA   Hoetzler G., Preisegger K.-H., Kurzbauer R., Denk H.;
RT   "Expression of three- and four-repeat tau isoforms in mouse liver.";
RL   Hepatology 20:1086-1089(1994).
RN   [3]
RP   SEQUENCE FROM N.A. (ISOFORMS TAU B AND TAU C).
RC   TISSUE=Brain;
RX   MEDLINE=89000516; PubMed=3123228;
RA   Lee G., Cowan N., Kirschner M.;
RT   "The primary structure and heterogeneity of tau protein from mouse
RT   brain.";
RL   Science 239:285-288(1988).
RN   [4]
RP   PARTIAL SEQUENCE FROM N.A. (ISOFORM B).
RC   STRAIN=BC; TISSUE=Brain;
RX   MEDLINE=95182802; PubMed=7877441;
RA   Sawa A., Oyama F., Matsushita M., Ihara Y.;
RT   "Molecular diversity at the carboxyl terminus of human and rat tau";
RL   Brain Res. Mol. Brain Res. 27:111-117(1994).
RN   [5]
RP   ...ASSOCIATION.
RX   MEDLINE=94009827; PubMed=8492762;
RA   Couchie D., Gache Y., Mavilia C., Guillemot J., Briand A.-M.,
RA   Nivez M.-P., Nunez J.;
RT   "High molecular weight tau proteins and acquisition of neuronal
RT   polarity in peripheral nervous system.";
RL   C. R. Acad. Sci., III, Sci. Vie 316:1043-1047(1993).
CC   -!- FUNCTION: PROMOTES MICROTUBULE ASSEMBLY AND STABILITY, AND MIGHT
CC       BE INVOLVED IN THE ESTABLISHMENT AND MAINTENANCE OF NEURONAL
CC       POLARITY. THE C-TERMINUS BINDS AXONAL MICROTUBULES WHILE THE N-
CC       TERMINUS BINDS NEURAL PLASMA MEMBRANE COMPONENTS, SUGGESTING THAT
CC       TAU FUNCTIONS AS A LINKER PROTEIN BETWEEN BOTH AXONAL F-ACTIN AND
CC       MICROTUBULES. DETERMINED BY TAU PHOSPHORYLATION (IN THE NEURONAL CELL) IN THE
CC       DOMAIN OF THE CELL BODY DEFINED BY THE CYTOSKELETON. THE SIX BRAIN
CC       ISOFORMS ALLOW PLASTICITY OF THE CYTOSKELETON WHEREAS THE LONGER
CC       ISOFORMS MAY PREFERENTIALLY PLAY A ROLE IN ITS STABILIZATION.
CC   -!- SUBCELLULAR LOCATION: MOSTLY FOUND IN THE AXONS OF NEURONS, IN THE
CC       CYTOSOL AND IN ASSOCIATION WITH PLASMA MEMBRANE COMPONENTS.
CC   -!- ALTERNATIVE PRODUCTS: AT LEAST 6 ISOFORMS; PNS TAU (SHOWN HERE),
CC       TAU-A, TAU-B, TAU-C, TAU-D AND TAU-E; ARE PRODUCED BY ALTERNATIVE
CC       SPLICING. THEY DIFFER FROM EACH OTHER BY THE PRESENCE OR ABSENCE
CC       OF UP TO 5 OF THE 14 EXONS. ONE OF THESE OPTIONAL EXONS CONTAINS
CC       THE ADDITIONAL TAU/MAP REPEAT. TWO DIFFERENT C-TERMINI ARE
CC       OBTAINED EITHER BY THE RETENTION OR THE SPLICING OF INTRON 13.
CC   -!- TISSUE SPECIFICITY: EXPRESSED IN NEURONS AND AT A LOWER LEVEL IN
CC       THE LIVER AND KIDNEY. PNS TAU IS EXPRESSED IN THE PERIPHERAL
CC       NERVOUS SYSTEM WHILE THE OTHERS ARE EXPRESSED IN THE CENTRAL
CC       NERVOUS SYSTEM.
CC   -!- DEVELOPMENTAL STAGE: SHORTER FORMS OF LOW MOLECULAR WEIGHT TAU
CC       (LMW-TAU) ARE GENERALLY EXPRESSED AT EARLY DEVELOPMENT STAGE, AND
CC       LONGER FORMS OF HIGH MOLECULAR WEIGHT TAU (HMW TAU) IN THE ADULT
CC       BRAIN.
CC   -!- DOMAIN: THE TAU/MAP REPEAT BINDS TO TUBULIN. TYPE I (3 OR 4 SUB-
CC       CONTAIN 3 REPEATS WHILE TYPE II (2 OR 3 SUB-CONTAIN 4 REPEATS.
CC   -!- PTM: PHOSPHORYLATION AT VARIOUS SERINE AND THREONINE RESIDUES.  MOST
CC       S-P OR T-P MOTIFS BY PROLINE DIRECTED PROTEIN KINASES (PEPK) (...,
CC       CDK5, GSK3, MAPK) (A FEW SITES PER PROTEIN IN INTERPHASE, MORE IN
CC       MITOSIS). AND AT SERINE RESIDUES IN K-X-G-S MOTIFS BY
CC       MAP/MICROTUBULE AFFINITY REGULATING KINASE (MARK) (BY SIMILARITY).
CC   -!- DISEASE: MAY BE INVOLVED IN THE TAU GENESIS I.E. NEUROFIBRILLARY
CC       INCLUSIONS (AS MAY BE SEEN IN DIVERSE AND MORE RECENTLY
CC       IDENTIFIED NEURODEGENERATIVE DISEASES (E.G. ALZHEIMER'S
CC       DIHYDROCOLLIDINE), A MODEL FOR HUMAN ALCOHOLIC HEPATITIS.
CC       ALTERATION OF TAU (ABNORMAL PHOSPHORYLATION AND CROSS-LINKING)
CC       COULD CONTRIBUTE TO MALE-SEX BODIES FORMATION AND INFERTILITY.
CC       MICROTUBULE FUNCTION IN ALCOHOLIC LIVER DISEASE.
CC   -!- SIMILARITY: CONTAINS 4 TAU/MAP REPEATS.
CC   ---------- ---------- ---------- ---------- ---------- ----------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed.  Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------- ---------- ---------- ---------- ---------- ----------
DR   EMBL; U12914; AAA58343.1; -.
DR   EMBL; U12915; AAA58344.1; -.
DR   EMBL; U12916; AAA58345.1; -.
DR   EMBL; Z12133; CAA78121.1; -.
DR   EMBL; M94266; -; NOT_ANNOTATED_CDS.
DR   EMBL; M18777; AAA40169.1; -.
DR   EMBL; M18776; AAA40166.1; -.
DR   EMBL; D30627; BAA18878.1; -.
DR   PIR; A28820; A28820.
DR   PIR; B28820; B28820.
DR   MGD; MGI:97180; Mapt.
DR   InterPro; IPR002955; Tau_protein.
DR   InterPro; IPR001084; Tubulin_bind.
DR   Pfam; PF00418; tubulin_binding; 7.
DR   PRINTS; PR01261; TAUPROTEIN.
```

```
DR   PROSITE; PS00229; TAU_MAP; 4.
KW   Microtubule; Cytoskeleton; Repeat; Alternative splicing; Acetylation;
KW   Phosphorylation.
FT   INIT_MET      1      1       BY SIMILARITY.
FT   REPEAT      535    565       TAU/MAP MOTIF 1.
FT   REPEAT      566    596       TAU/MAP MOTIF 2.
FT   REPEAT      597    627       TAU/MAP MOTIF 3.
FT   REPEAT      628    659       TAU/MAP MOTIF 4.
FT   MOD_RES       1      1       ACETYLATION (BY SIMILARITY).
FT   DISULFID    582    615       BY SIMILARITY.
FT   VARSPLIC     33     98       MISSING (IN ISOFORM TAU B, ISOFORM TAU C,
FT                                ISOFORM TAU D AND ISOFORM TAU E).
FT   VARSPLIC     91     11.      MISSING (IN ISOFORM TAU E).
FT   VARSPLIC    113     34       MISSING (IN ISOFORM TAU A, ISOFORM TAU B,
FT                                ISOFORM TAU C, ISOFORM TAU D AND ISOFORM
FT                                TAU E).
FT   VARSPLIC    367     43.      MISSING (IN ISOFORM TAU A, ISOFORM TAU B,
FT                                ISOFORM TAU C, ISOFORM TAU D AND ISOFORM
FT                                TAU E).
FT   VARSPLIC    566     59       MISSING (IN ISOFORM TAU B AND ISOFORM
FT                                TAU C).
FT   VARSPLIC    732     73.      L   -> KAAEESQVWNYSHLAFTHSGL (IN
FT                                ISOFORM TAU B).
FT   CONFLICT      2       .      D   -> N (IN REF. 1).
FT   CONFLICT      8       .      D   -> N (IN REF. 1).
FT   CONFLICT    527     527      P   -> T (IN REF. 2; CAA78121).
FT   CONFLICT    671     671      E   -> Q (IN REF. 1).
SQ   SEQUENCE   732 AA  76112 MW;  BEDF0767B41C7A3A CRC64;


  Query Match             99.8%;  Score 733;  DB 1;  Length 732;
  Best Local Similarity   99.6%;  Pred. No. 99;
  Matches  727;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 FSKIASNIQQ 10
           ||||||| ||
Db     248 FSKVSAEIQQ 257


RESULT  15
TAU_RAT
ID   TAU_RAT        STANDARD;      PRT;   732 AA.
AC   P19332; Q63567; Q60WD0; Q63677;
DT   01-NOV-1990 (Rel. 16, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
DE   Microtubule-associated protein tau (Neurofibrillary tangle protein)
DE   (Paired helical filament-tau) (PHF-tau).
GN   MAPT OR MTBT OR TAU.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORM TAU A).
RC   TISSUE=Pheochromocytoma;
RX   MEDLINE=92179685; PubMed=1532...;
RA   Goedert M., Spillantini M.G., Crowther R.A.;
RT   "Cloning of a big tau microtubule-associated protein characteristic of
RT   the peripheral nervous system.";
RL   Proc. Natl. Acad. Sci. U.S.A. ...(1992).
RN   [2]
RP   SEQUENCE FROM N.A. (ISOFORM TAU B).
RC   TISSUE=Dorsal root ganglion;
RX   MEDLINE=94173081; PubMed=84...;
RA   Georgieff I.S., Liem R.K.H., Couchie D., Mavilia C., Nunez J.,
RA   Shelanski M.L.;
RT   "Expression of high molecular weight tau in the central and peripheral
RT   nervous systems.";
RL   J. Cell Sci. 105:729-737(1993).
RN   [3]
RP   SEQUENCE FROM N.A. (ISOFORM TAU F).
RC   STRAIN=Wistar; TISSUE=Brain;
RX   MEDLINE=94334997; PubMed=8057376;
RA   Sadot E., Marx R., Barg J., Behar L., Ginzburg I.;
RT   "Complete sequence of 3'-untranslated region of tau from rat central
RT   nervous system. Implications for mRNA heterogeneity.";
RL   J. Mol. Biol. 241:325-331(1994).
RN   [4]
RP   SEQUENCE FROM N.A. (ISOFORMS TAU E AND TAU G).
RC   TISSUE=Brain;
RX   MEDLINE=90180457; PubMed=2561640;
RA   Kosik K.S., Orecchio L.D., Bakalis S., Neve R.L.;
RT   "Developmentally regulated expression of specific tau sequences.";
RL   Neuron 2:1389-1397(1989).
RN   [5]
RP   SEQUENCE FROM N.A. (ISOFORMS TAU E AND TAU C).
RX   MEDLINE=89295099; PubMed=2504728;
RA   Kanai Y., Takemura R., Oshima T., Mori H., Ihara Y., Yanagisawa M.,
RA   Masaki T., Hirokawa N.;
RT   "Expression of multiple tau isoforms and microtubule bundle formation
RT   in fibroblasts transfected with a single tau cDNA.";
RL   J. Cell Biol. 109:1173-1184(1989).
RN   [6]
RP   SEQUENCE OF 459-464 FROM N.A. (ISOFORM TAU A), AND SEQUENCE OF 1-5...
RP   AND 467-460 FROM N.A. (ISOFORM TAU D).
RC   TISSUE=Spinal cord;
RX   MEDLINE=95054048; PubMed=7964751;
RA   Mavilia C., Couchie D., Nunez J.;
RT   "Diversity of high molecular-weight tau proteins in different regions
RT   of the nervous system.";
RL   J. Neurochem. 63:2300-2306(1994).
RN   [7]
RP   SEQUENCE OF 696-751 FROM N.A. (ISOFORMS TAU A TO TAU G), AND SEQUENCE
RP   OF 751-774 FROM N.A. (ISOFORM TAU H).
RC   STRAIN=Sprague Dawley; TISSUE=Brain;
RX   MEDLINE=95182802; PubMed=7877441;
RA   Sawa A., Oyama F., Matsushita M., Ihara Y.;
RT   "Molecular diversity at the carboxyl terminus of human and rat tau.";
RL   Brain Res. Mol. Brain Res. 27:111-117(1994).
CC   -!- FUNCTION: PROMOTES MICROTUBULE ASSEMBLY AND STABILITY, AND MIGHT
CC       BE INVOLVED IN THE ESTABLISHMENT AND MAINTENANCE OF NEURONAL
CC       POLARITY. THE C-TERMINUS BINDS AXONAL MICROTUBULES WHILE THE N-
CC       TERMINUS BINDS NEURAL PLASMA MEMBRANE COMPONENTS, SUGGESTING THAT
CC       TAU FUNCTIONS AS A LINKER PROTEIN BETWEEN BOTH. AXONAL POLARITY IS
CC       PREDETERMINED BY TAU LOCALIZATION (IN THE NEURONAL CELL) IN THE
CC       DOMAIN OF THE CELL BODY DEFINED BY THE CENTROSOME. THE GROWTH
CC       ISOFORMS ALLOW PLASTICITY OF THE CYTOSKELETON WHEREAS THE LONGER
CC       ISOFORMS MAY PREFERENTIALLY PLAY A ROLE IN ITS STABILIZATION.
CC     ! SUBCELLULAR LOCATION: MOSTLY FOUND IN THE AXONS OF NEURONS. IN THE
CC       CYTOSOL AND IN ASSOCIATION WITH PLASMA MEMBRANE COMPONENTS.
CC     ! ALTERNATIVE PRODUCTS: AT LEAST 8 ISOFORMS, TAU A/671 (SHOWN HERE),
CC       TAU B/695 TAU/HMW TAU, TAU C, TAU D/SC2, TAU-E, TAU-F, TAU-
CC       G/FETAL TAU AND TAU H, ARE PRODUCED BY ALTERNATIVE SPLICING. THEY
CC       DIFFER FROM EACH OTHER BY THE PRESENCE OR ABSENCE OF UP TO 4 OF
CC       THE 14 EXONS. TWO DIFFERENT C-TERMINI ARE OBTAINED EITHER BY THE
CC       RETENTION OR THE SPLICING OF INTRON 13/14.
CC     ! TISSUE SPECIFICITY: EXPRESSED IN NEURONS. THE LARGER FORMS (TAU A
CC       AND TAU B) ARE PREFERENTIALLY EXPRESSED IN THE PERIPHERAL NERVOUS
CC       SYSTEM WHILE THE SHORTER ARE EXPRESSED IN THE CENTRAL NERVOUS
CC       SYSTEM. LOW AMOUNTS OF THE LARGER FORMS ARE ALSO FOUND IN LIMITED
CC       AREAS OF THE CNS.
CC     -!- DEVELOPMENTAL STAGE: DURING THE IMMEDIATE POSTNATAL PERIOD, THE
CC       NEURAL NETWORK EXPRESS ALL ISOFORMS WHEREAS ONLY THE LARGER
CC       FORMS PERSIST IN THE ADULTS.
CC     ! DOMAIN: THE TAU/MAP REPEAT BINDS TO TUBULIN. TYPE I ISOFORMS
CC       CONTAIN 3 REPEATS WHILE TYPE II ISOFORMS CONTAIN 4 REPEATS.
CC     ! PTM: PHOSPHORYLATION AT VARIOUS SERINE AND THREONINE RESIDUES IN
CC       S-P OR T-P MOTIFS BY PROLINE DIRECTED PROTEIN KINASES (PDPK):
CC       CDC2, GSK3, MAPK) (ALLOW BINDING FOR PROLYL ISOMERASE PIN1 IN
CC       MITOSIS), AND AT SERINE RESIDUES IN K-X-G-S MOTIFS BY
CC       MAP/MICROTUBULE AFFINITY REGULATING KINASE (MARK) (BY SIMILARITY).
CC     -!- SIMILARITY: CONTAINS 4 TAU/MAP REPEATS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation
```

CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

DR EMBL; M84175; AAA42204.1; -.
DR EMBL; X79321; CAA55889.1; -.
DR EMBL; D43628; -; NOT ANNOTATED CDS.
DR EMBL; D43629; -; NOT ANNOTATED CDS.
DR PIR; JS0406; JS0406.
DR PIR; A35574; A35574.
DR InterPro; IPR003095; Tau protein.
DR InterPro; IPR000484; Tubulin bind.
DR Pfam; PF00418; tubulin binding; 8.
DR PRINTS; PR01261; TAUPROTEIN.
DR PROSITE; PS00229; TAU_MAP; 4.
KW Microtubules; Cytoskeleton; Repeat; Alternative splicing; Acetylation;
KW Phosphorylation.
FT INIT_MET 0 0 BY SIMILARITY.
FT REPEAT 554 584 TAU/MAP MOTIF 1.
FT REPEAT 585 615 TAU/MAP MOTIF 2.
FT REPEAT 616 646 TAU/MAP MOTIF 3.
FT REPEAT 647 678 TAU/MAP MOTIF 4.
FT MOD_RES 1 1 ACETYLATION (BY SIMILARITY).
FT DISULFID 671 832 BY SIMILARITY.
FT VARSPLIC 33 90 MISSING (IN ISOFORM TAU F).
FT VARSPLIC 62 90 MISSING (IN ISOFORM TAU F AND
FT ISOFORM TAU G).
FT VARSPLIC 113 369 MISSING (IN ISOFORM TAU C, ISOFORM TAU D,
FT ISOFORM TAU E, ISOFORM TAU F AND ISOFORM
FT TAU G).
FT VARSPLIC 369 401 MISSING (IN ISOFORM TAU B, ISOFORM TAU C,
FT ISOFORM TAU E, ISOFORM TAU F AND ISOFORM
FT TAU G).
FT VARSPLIC 595 615 MISSING (IN ISOFORM TAU G).
FT VARSPLIC 291 756 L -> KPVLISSEVWNYSHLPSHHTSLSS (IN ISOFORM
FT TAU H).
FT CONFLICT 254 254 F -> L (IN REF. 2).
FT CONFLICT 283 283 G -> A (IN REF. 2).
FT CONFLICT 291 291 H -> D (IN REF. 2).
FT CONFLICT 617 617 H -> Q (IN REF. 2, 3 AND 4).
FT CONFLICT 734 734 Y -> H (IN REF. 3).
FT CONFLICT 733 733 P -> A (IN REF. 2).
SQ SEQUENCE 751 AA; 78432 MW; B96B732B4441B4B2 CRC64;

Query Match 99.8%; Score 36; DB 1; Length 751;
Best Local Similarity 55.6%; Pred. No. 55;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 PSKIASNI_ 9
|||::: ||
DB 267 PSKVSAETQ 275

Search completed: August 26, 2002, 06:29:04
Job time: 960 sec

SmithWaterman version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

M protein - protein search, using sw model

Run on: August 26, 2002, 06:28:07 ; Search time 130.19 Seconds
(without alignments)
12.959 Million cell updates/sec

Title: US-09-744-282-4
Perfect score: 43
Sequence: 1 FSKIASNIQ 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172694929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_19:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_primate:*
16: sp_bacteriaps:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	43	100.0	383	4	Q96H17	Q96H17 homo sapien
2	43	100.0	383	6	Q29411	Q29411 sus scrofa
3	37	86.0	332	6	Q18949	Q18949 bos taurus
4	36	83.7	381	11	Q99J84	Q99J84 mus musculu
5	35	81.4	62	5	Q6QV34	Q6QV34 hydra magni
6	33	76.7	576	4	Q9Y3V7	Q9Y3V7 homo sapien
7	33	76.7	888	6	Q97867	Q97867 sus scrofa
8	33	76.7	1042	4	Q13792	Q13792 homo sapien
9	33	76.7	1081	4	Q76065	Q76065 homo sapien
10	32	74.4	151	2	Q7AEB0	Q7AEB0 lactobacill
11	32	74.4	804	5	Q9NA22	Q9NA22 caenorhabdi
12	32	74.4	2229	5	Q19853	Q19853 caenorhabdi
13	31	72.1	107	11	Q9D648	Q9D648 mus musculu
14	31	72.1	140	5	Q6V9A1	Q6V9A1 drosophila
15	31	72.1	279	16	Q92EE9	Q92EE9 listeria in
16	31	72.1	333	2	Q54065	Q54065 salmonella
17	31	72.1	335	2	Q57177	Q57177 salmonella
18	31	72.1	336	2	Q54044	Q54044 salmonella
19	31	72.1	336	2	Q54046	Q54046 salmonella
20	31	72.1	336	2	Q54048	Q54048 salmonella
21	31	72.1	336	2	Q54050	Q54050 salmonella
22	31	72.1	336	2	Q54052	Q54052 salmonella
23	31	72.1	336	2	Q54054	Q54054 salmonella
24	31	72.1	336	2	Q54056	Q54056 salmonella
25	31	72.1	336	2	Q54057	Q54057 salmonella
26	31	72.1	336	2	Q54061	Q54061 salmonella
27	31	72.1	343	2	Q54059	Q54059 salmonella
28	31	72.1	343	2	Q54060	Q54060 salmonella
29	31	72.1	347	4	Q99PK6	Q99PK6 homo sapien
30	31	72.1	357	9	Q94MK7	Q94MK7 bacteriopha
31	31	72.1	357	9	Q94MK2	Q94MK2 bacteriopha
32	31	72.1	357	9	Q94MJ7	Q94MJ7 bacteriopha
33	31	72.1	357	9	Q94MI8	Q94MI8 bacteriopha
34	31	72.1	357	9	Q94MI3	Q94MI3 bacteriopha
35	31	72.1	357	9	Q94MH9	Q94MH9 bacteriopha
36	31	72.1	357	9	Q94MH4	Q94MH4 bacteriopha
37	31	72.1	357	9	Q94MG9	Q94MG9 bacteriopha
38	31	72.1	357	9	Q94MF9	Q94MF9 bacteriopha
39	31	72.1	357	9	Q94MF4	Q94MF4 bacteriopha
40	31	72.1	357	9	Q94MF0	Q94MF0 bacteriopha
41	31	72.1	357	9	Q94ME5	Q94ME5 bacteriopha
42	31	72.1	357	9	Q94ME0	Q94ME0 bacteriopha
43	31	72.1	357	9	Q94MD1	Q94MD1 bacteriopha
44	31	72.1	357	9	Q94MC7	Q94MC7 bacteriopha
45	31	72.1	357	9	Q941Z4	Q941Z4 bacteriopha

ALIGNMENTS

RESULT 1
Q96H17
ID Q96H17 PRELIMINARY; PRT; 383 AA.
AC Q96H17;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE SIMILAR TO CHITINASE 3-LIKE 1 (CARTILAGE GLYCOPROTEIN-39).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BRAIN, AND GLIOBLASTOMA WITH EGFR AMPLIFICATION;
RA Strausberg R.;
RL Submitted (MAY 2001) to the EMBL/GenBank/DDBJ databases.
DR EMBL; BC008568; AAH08568.1; -.
SQ SEQUENCE 383 AA; 42625 MW; 6C9EF133B0C7C2D1 CRC64;

Query Match 100.0%; Score 43; DB 4; Length 383;
Best Local Similarity 100.0%; Pred. No. 0.26;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FSKIASNIQ 9
||||||||| |
Db 106 FSKIASNIQ 114

RESULT 2
Q29411
ID Q29411 PRELIMINARY; PRT; 383 AA.
AC Q29411;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE 38 KDA HEPARIN BINDING GLYCOPROTEIN.
OS Sus scrofa (Pig).

```
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX   NCBI_TaxID=9823;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=SMOOTH MUSCLE;
RX   MEDLINE=95305899; PubMed=7768902;
RA   Shackelton L.M., Mann D.M., Millis A.J.;
RT   "Identification of a 38 kDa heparin-binding glycoprotein (gp38k) in
RT   differentiating vascular smooth muscle cells as a member of a group of
RT   proteins associated with tissue remodeling.";
RL   J. Biol. Chem. 270:13076-13083(1995).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=SMOOTH MUSCLE;
RA   Shackelton L.M., Mann D.M., Millis A.J.;
RL   Submitted (JAN-1996) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U19900; AAA86482.1; -.
DR   EMBL; Z47803; CAA87764.1; -.
DR   InterPro; IPR001223; Glyco_hydro_18.
DR   Pfam; PF00704; Glyco_hydro_18; 1.
KW   Heparin-binding.
SQ   SEQUENCE   383 AA;  42443 MW;  3D1039F499108DC4 CRC64;


  Query Match          100.0%;  Score 43;  DB 6;  Length 383;
  Best Local Similarity  100.0%;  Pred. No. 0.26;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 FSKIASNIQ 9
           |||||||||
Db     106 FSKIASNIQ 114


RESULT   3
O18949
ID   O18949       PRELIMINARY;      PRT,   332 AA.
AC   O18949;
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   CHITINASE LIKE PROTEIN 1 (FRAGMENT).
GN   CLP-1.
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Recklies A.D., White C.;
RT   "Expression of chitinase-like protein 1 (CLP-1) in bovine
RT   chondrocytes.";
RL   Submitted (JUN-1997) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF011373; AAB64304.1; -.
DR   InterPro; IPR001223; Glyco_hydro_18.
DR   Pfam; PF00704; Glyco_hydro_18; 1.
FT   NON_TER       1       1
FT   NON_TER     332     332
SQ   SEQUENCE   332 AA;  37434 MW;  03E163234864D8E CRC64;


  Query Match          86.0%;  Score 37;  DB 6;  Length 332;
  Best Local Similarity  88.9%;  Pred. No. 4.8;
  Matches    8;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 FSKIASNIQ 9
           |||| ||||
Db      82 FSKIASKIQ 90


RESULT   4
Q9CIP4
ID   Q99184       PRELIMINARY;      PRT,   381 AA.
AC   Q99184;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-OCT-2001 (TrEMBLrel. 18, Last annotation update)
DE   SIMILAR TO CHITINASE 3 LIKE 1 (CARTILAGE GLYCOPROTEIN 39).
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Straunberg R.;
RL   Submitted (AUG-2001) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Straunberg R.;
RL   Submitted (FEB-2001) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RA   Straunberg R.;
RL   Submitted (MAR-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC005611; AAH05611.1; -.
DR   EMBL; BC003780; AAH03780.1; -.
DR   EMBL; BC004734; AAH04734.1; -.
DR   InterPro; IPR001223; Glyco_hydro_18.
DR   Pfam; PF00704; Glyco_hydro_18; 1.
SQ   SEQUENCE   381 AA;  42979 MW;  EE6588C9AB9D445D CRC64;


  Query Match          88.7%;  Score 39;  DB 6;  Length 381;
  Best Local Similarity  77.8%;  Pred. No. 9.2;
  Matches    7;  Conservative    2;  Mismatches    0;  Indels    1;  Gaps    1;

Qy       1 FSKIASNIQ 9
           ||:||||:|
Db     107 FSEIASNTE 115


RESULT   5
Q9GV84
ID   Q9GV84       PRELIMINARY;      PRT,    62 AA.
AC   Q9GV84;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   HYM-323 PROTEIN.
GN   HYM-323.
OS   Hydra magnipapillata (Hydra).
OC   Eukaryota; Metazoa; Cnidaria; Hydrozoa; Hydroida; Anthoathecatae;
OC   Hydridae; Hydra.
OX   NCBI_TaxID=6085;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=105;
RX   MEDLINE=21067881; PubMed=11152642;
RA   Hatayuji N., Takahashi T., Hatta M., Tezuka H., Morishita T.;
RA   Matsushima O., Fujisawa T.;
RT   "Enhancement of foot formation in Hydra by a novel peptide, peptide
RT   Hym-323.";
RL   Development 128:437-446(2001).
DR   EMBL; AB040074; BAB13535.1; -.
SQ   SEQUENCE    62 AA;  7162 MW;  A04A9220C90E7DDF CRC64;


  Query Match          81.4%;  Score 35;  DB 6;  Length 62;
  Best Local Similarity  87.5%;  Pred. No. 2.4;
  Matches    7;  Conservative    1;  Mismatches    0;  Indels    1;  Gaps    1;

Qy       1 FSKIASNT 8
           |:||||||
Db      15 FSEIASNE 22
```

RESULT 6
Q9Y4V7
ID Q9Y4V7 PRELIMINARY; PRT; 576 AA.
AC Q9Y4V7;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE HYPOTHETICAL 63.5 KDA PROTEIN (FRAGMENT).
GN DKFZp586A1524.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=EMBRYO;
RA Wiemann S., Bechtel S., Mewes H.W., Sassenhagen J., Weinstock G.;
RL Submitted (MAR-1999) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AL136095; CAB43267.1; -.
DR HSSP; P00736; 1AL1.
DR InterPro; IPR000152; Asp_Lys_Asn_HS.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR001881; EGF_Ca.
DR Pfam; PF00008; EGF; 6.
DR SMART; SM00179; EGF_CA; 8.
DR SMART; SM00181; EGF_like; 2.
DR PROSITE; PS00010; ASX_HYDROXYL; 4.
DR PROSITE; PS01186; EGF_2; 4.
DR PROSITE; PS01187; EGF_CA; 9.
KW Calcium binding; EGF-like domain; Glycoprotein; Hydroxylation;
KW Hypothetical protein; Repeat.
FT NON_TER 1 1
SQ SEQUENCE 576 AA; 64274 MW; F12E3E9381C18A61 CRC64;

Query Match 76.7%; Score 33; DB 4; Length 576;
Best Local Similarity 66.0%; Pred. No. 64;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FSKIASNIQ 9
|||:|:|||
Db 210 FSQVASNIT 218

RESULT 7
Q76B67
ID Q76B67 PRELIMINARY; PRT; 988 AA.
AC Q76B67;
DT 01-MAY-2003 (TrEMBLrel. 21, Created)
DT 01-MAY-2003 (TrEMBLrel. 21, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE GASTRIC MUCIN (FRAGMENT).
GN MUC5AC.
OS Sus scrofa domesticus (domestic pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9825;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=GASTRIC EPITHELIUM;
RX MEDLINE=95293957; PubMed=7775418;
RA Desseyn J.L., Buisine M.P., Porchet N., Aubert J.P., Laine A.;
RA Zweibaum A., Real F.X.;
RT "Characterisation of a mucin cDNA clone isolated from HT-29 mucus
RT secreting cells: the 3' end of MUC5AC2";
RL J. Biol. Chem. 270:13665-13673(1995).
DR EMBL; Z48314; CAA88307.1; -.
DR InterPro; IPR000359; Cys_knot.
DR InterPro; IPR002405; GF_cysknot.
DR InterPro; IPR002919; TIL.
DR InterPro; IPR001846; Vwd.
DR InterPro; IPR001007; VWFC.
DR Pfam; PF00007; Cys_knot; 1.
DR Pfam; PF01826; TIL; 1.
DR Pfam; PF00093; vwc; 1.
DR Pfam; PF00094; vwd; 1.
DR PRINTS; PR00438; GFCYSKNOT.
DR SMART; SM00041; CT; 1.
DR SMART; SM00214; VWC; 3.
DR SMART; SM00216; VWD; 1.
DR PROSITE; PS01185; CTCK_1; 1.
DR PROSITE; PS01225; CTCK_2; 1.
DR PROSITE; PS01208; VWFC; 2.
FT NON_TER 1 1

RT factor and cystine knot domains at the carboxyl terminal(1).";
RL Biochim. Biophys. Acta 1447:77-92(1999).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=GASTRIC EPITHELIUM;
RA Turner B.S., Bhaskar K.R., Hadzopoulou-Cladaras M., LaMont J.T.;
RT Submitted (MAR-1998) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AF054584; AAD19833.1; -.
DR InterPro; IPR000359; Cys_knot.
DR InterPro; IPR002405; GF_cysknot.
DR InterPro; IPR002919; TIL.
DR InterPro; IPR001846; Vwd.
DR InterPro; IPR001007; VWFC.
DR Pfam; PF00007; Cys_knot; 1.
DR Pfam; PF01826; TIL; 1.
DR Pfam; PF00093; vwc; 1.
DR Pfam; PF00094; vwd; 1.
DR PRINTS; PR00438; GFCYSKNOT.
DR SMART; SM00041; CT; 1.
DR SMART; SM00214; VWC; 2.
DR SMART; SM00216; VWD; 1.
DR PROSITE; PS01185; CTCK_1; 1.
DR PROSITE; PS01225; CTCK_2; 1.
DR PROSITE; PS01208; VWFC; 2.
FT NON_TER 1 1
SQ SEQUENCE 988 AA; 105264 MW; 786F949E0FCE8465 CRC64;

Query Match 76.7%; Score 33; DB 6; Length 988;
Best Local Similarity 66.7%; Pred. No. 1.1e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FSKIASNIQ 9
|||:|:|||
Db 410 FSKFANNIE 418

RESULT 8
Q13792
ID Q13792 PRELIMINARY; PRT; 1042 AA.
AC Q13792;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE APOMUCIN (FRAGMENT).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95293957; PubMed=7775418;
RA Lesuffleur T., Kornowski R., Barbat A., Bara J., Fleiszer M., Bara M., Swallow D.M.,
RA Zweibaum A., Real F.X.;
RT "Characterisation of a mucin cDNA clone isolated from HT-29 mucus
RT secreting cells: the 3' end of MUC5AC2";
RL J. Biol. Chem. 270:13665-13673(1995).
DR EMBL; Z48314; CAA88307.1; -.
DR InterPro; IPR000359; Cys_knot.
DR InterPro; IPR002405; GF_cysknot.
DR InterPro; IPR001846; Vwd.
DR InterPro; IPR001007; VWFC.
DR Pfam; PF00007; Cys_knot; 1.
DR Pfam; PF00093; vwc; 1.
DR Pfam; PF00094; vwd; 1.
DR PRINTS; PR00438; GFCYSKNOT.
DR SMART; SM00041; CT; 1.
DR SMART; SM00214; VWC; 3.
DR SMART; SM00216; VWD; 1.
DR PROSITE; PS01185; CTCK_1; 1.
DR PROSITE; PS01225; CTCK_2; 1.
DR PROSITE; PS01208; VWFC; 2.
FT NON_TER 1 1

```
SQ   SEQUENCE   1042 AA;  111189 MW;  E17801C2B83C00C2 CRC64;


  Query Match             76.7%;  Score 33;  DB 4;  Length 1042;
  Best Local Similarity   66.7%;  Pred. No. 1.1e+02;
  Matches    6;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 FSKIASNQ 9
             ||| |:||
Db     348 FSKEANNIE 356


RESULT   9
Q76065
ID   Q76065       PRELIMINARY;      PRT;  1081 AA.
AC   Q76065;
DT   01-NOV-1998 (TrEMBLrel. 08, Created)
DT   01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   MUC5AC PROTEIN (FRAGMENT).
GN   MUC5AC.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=TRACHEA, AND PLACENTA;
RX   MEDLINE=98285941; PubMed=9620876;
RA   Buisine M.P., Desseyn J.L., Porchet N., Degand P., Laine A.,
RA   Aubert J.P.;
RT   "Genomic organization of the 3' region of the human MUC5AC mucin gene:
RT   additional evidence for a common ancestral gene for the cysteine-rich
RT   gene family.";
RL   Biochem. J. 332:729-738(1998).
DR   EMBL; AJ001402; CAA04737.1; -.
DR   EMBL; AJ001403; CAA04738.1; -.
DR   InterPro; IPR000859; Cys_knot.
DR   InterPro; IPR002400; GF_cysknot.
DR   InterPro; IPR001846; Vwd.
DR   InterPro; IPR001007; VWFC.
DR   Pfam; PF00007; Cys_knot; 1.
DR   Pfam; PF00094; vwc; 1.
DR   Pfam; PF00094; vwd; 1.
DR   PRINTS; PR00438; GFCYSKNOT.
DR   SMART; SM00041; CT; 1.
DR   SMART; SM00214; VWC; 3.
DR   SMART; SM00216; VWD; 1.
DR   PROSITE; PS01185; CTCK_1; 1.
DR   PROSITE; PS01225; CTCK_2; 1.
DR   PROSITE; PS51208; VWFC; 2.
FT   NON_TER       1       1
SQ   SEQUENCE   1081 AA;  115988 MW;  A40E779548650A5C CRC64;


  Query Match             76.7%;  Score 33;  DB 4;  Length 1081;
  Best Local Similarity   66.7%;  Pred. No. 1.2e+02;
  Matches    6;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 FSKIASNTQ 9
             ||| |:||
Db     479 FSKEANNIE 487


RESULT  10
Q9AED0
ID   Q9AED0       PRELIMINARY;      PRT;   151 AA.
AC   Q9AED0;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   HYPOTHETICAL 17.6 KDA PROTEIN.
GN   INP1516DEC.
OS   Lactobacillus sakei 14'.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales;
OC   Lactobacillus.
OX   NCBI_TaxID=1644;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=145;
RA   Skaugen M.;
RT   "Identification, characterization and expression of a novel,
RT   bicistronic operon involved in the lactocin S production of
RT   Lactobacillus sakei 145.";
RL   Submitted (FEB-2001) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=145;
RX   MEDLINE=94368096; PubMed=7521995;
RA   Skaugen M.;
RT   "Transposition in Lactobacillus sakei and its mediation of a novel
RT   protection by insertion of IS1163, a new member of the IS3 family.";
RL   Appl. Environ. Microbiol. 60:2808-2815(1994).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=145;
RX   MEDLINE=97234604; PubMed=9072078;
RA   Skaugen M.;
RT   "Organization and expression of a gene cluster involved in the
RT   biosynthesis of the antilisterial bacteriocin lactocin S.";
RL   Mol. Gen. Genet. 253:674-686(1997).
RN   [4]
RP   SEQUENCE FROM N.A.
RC   STRAIN=145;
RX   MEDLINE=20290259; PubMed=10832642;
RA   Skaugen M.;
RT   "Transposition in Lactobacillus sakei: insertion of a novel
RT   lactocin S operon by the insertion of IS1520, a new member of the IS3
RT   family of insertion sequences.";
RL   Microbiology 146:1351-1359(2000).
RN   [5]
RP   SEQUENCE FROM N.A.
RC   STRAIN=145;
RX   MEDLINE=95050447; PubMed=7961427;
RA   Skaugen M.;
RT   "In vivo conversion of L-Serine to D-Alanine in a ribosomally
RT   synthesized polypeptide.";
RL   J. Biol. Chem. 269:27183-27188(1994).
DR   EMBL; Z54312; CAA83838.1; -.
DR   InterPro; IPR000445; Cyt_c_heme_bind.
DR   PROSITE; PS00190; CYTOCHROME_C; UNKNOWN 1.
KW   Hypothetical protein.
SQ   SEQUENCE   151 AA;  16665 MW;  7F0E4AAE27D0AA4C CRC64;


  Query Match             74.4%;  Score 32;  DB 2;  Length 151;
  Best Local Similarity   75.0%;  Pred. No. 28;
  Matches    6;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 FSKIASNI 8
             ||| ||||
Db      66 FSKVLSNI 73


RESULT  11
Q9NA22
ID   Q9NA22       PRELIMINARY;      PRT;   834 AA.
AC   Q9NA22;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   ZK520.4A PROTEIN.
GN   ZK520.4A.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditina;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
```

```
OX   NCBI_TaxID=6239;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Steward C.A.;
RL   Submitted (MAR-1997) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=99006014; PubMed=9851916;
RA   none;
RT   "Genome sequence of the nematode C.elegans: A platform for
RT   investigating biology.";
RL   Science 282:2011-2012(1998).
DR   EMBL; Z92822; CAB07088.1; -.
DR   InterPro; IPR001478; PDZ.
DR   Pfam; PF00595; PDZ; 2.
DR   SMART; SM00182; PDZ; 1.
DR   PROSITE; PS50106; PDZ; 1.
DR   PROSITE; PS00660; PDZ_1; 1.
SQ   SEQUENCE   894 AA;  98281 MW;  86C1CB9E547E146E CRC64;


  Query Match             74.4%;  Score 32;  DB 5;  Length 894;
  Best Local Similarity   75.0%;  Pred. No. 1.4e+02;
  Matches    6;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2 SKIASNTQ 9
            :|:||||
Db     553 AKLASNIQ 560


RESULT  12
Q19858
ID   Q19858       PRELIMINARY;      PRT;  2229 AA.
AC   Q19858;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   31-DEC-2001 (TrEMBLrel. 16, Last annotation update)
DE   HYPOTHETICAL 244.9 KDA PROTEIN.
GN   F28B4.3.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BRISTOL N2;
RX   MEDLINE=99006014; PubMed=9851916;
RA   None.;
RT   "Genome sequence of the nematode C. elegans: a platform for
RT   investigating biology. The C. elegans Sequencing Consortium.";
RL   Science 282:2012-2018(1998).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BRISTOL N2;
RA   Coulson A.;
RT   "The sequence of C. elegans cosmid F28B4.";
RL   Submitted (JAN-1996) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BRISTOL N2;
RA   Waterston R.;
RT   "Direct Submission.";
RL   Submitted (JUN-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; Z42834; AAA83784.1; -.
DR   InterPro; IPR001478; PDZ; 4.
DR   InterPro; IPR001881; EGF_Ca; 1.
DR   InterPro; IPR001791; laminin_G; 1.
DR   InterPro; IPR002035; vWFA; 1.
DR   Pfam; PF00008; EGF; 2.
DR   Pfam; PF00054; laminin_G; 1.
DR   Pfam; PF00092; vwa; 1.
DR   SMART; SM00179; EGF_CA; 1.
DR   SMART; SM00181; EGF; 3.
DR   SMART; SM00327; VWA; 1.
DR   PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
DR   PROSITE; PS00022; EGF_1; UNKNOWN_6.
DR   PROSITE; PS01186; EGF_2; 4.
DR   PROSITE; PS50234; VWFA; 1.
KW   EGF-like domain; Glycoprotein; Hypothetical protein.
SQ   SEQUENCE   2229 AA;  244874 MW;  87AC82B6A97AEA25 CRC64;


  Query Match             71.4%;  Score 32;  DB 5;  Length 2229;
  Best Local Similarity   66.7%;  Pred. No. 4.9e+02;
  Matches    6;  Conservative    2;  Mismatches    1;  Indels    1;  Gaps    1;

Qy       1 FSKIASNTQ 9
             :| |:|||
Db    1231 YSSIAANIQ 1239


RESULT  13
Q9D648
ID   Q9D648       PRELIMINARY;      PRT;  107 AA.
AC   Q9D648;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   4731417B20RIK PROTEIN.
GN   4731417B20RIK.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciuromorpha; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=BRAIN;
RX   MEDLINE=21085660; PubMed=11217851;
RA   Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA   Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA   Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA   Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA   Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA   Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA   Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA   Schriml L.M., Staniere F., Suzuki R., Tomita M., Wagner L., Wasserman W.,
RA   Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA   Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA   Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA   Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA   Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA   Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA   Sasaki H., Sato K., Schonbach C., Seya T., Shibata Y., Storch K.-F.,
RA   Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA   Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA   Hayashizaki Y.;
RT   "Functional annotation of a full-length mouse cDNA collection.";
RL   Nature 409:685-690(2001).
CC   EMBL; AK014062, BAB29472.1; -.
DR   MGD; MGI:1921826; 4731417B20Rik.
SQ   SEQUENCE   107 AA;  12039 MW;  48210019E69C59F9 CRC64;


  Query Match             72.1%;  Score 31;  DB 11;  Length 107;
  Best Local Similarity   75.0%;  Pred. No. 88;
  Matches    6;  Conservative    2;  Mismatches    0;  Indels    1;  Gaps    1;

Qy       2 SKIASNTQ 9
            :|:||:|
Db     49 SKLASNSQ 56


RESULT  14
Q9V8A1
ID   Q9V8A1       PRELIMINARY;      PRT;  146 AA.
```

```
AC   Q9V8A1;
DT   31 MAY 2000 (TrEMBLrel. 13, Created)
DT   01 MAY 2000 (TrEMBLrel. 13, Last sequence update)
DT   01 DEC 2001 (TrEMBLrel. 19, Last annotation update)
DE   CG12567 PROTEIN.
GN   CG12567.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN BERKELEY;
RX   MEDLINE=20796039; PubMed=10731132;
RA   Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA   Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA   George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA   Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA   Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA   Wan K.H., Doyle C., Baxter E.G., Holt G., Nelson C.R., Miklos G.L.G.,
RA   Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA   Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA   Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA   Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA   Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA   Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA   de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA   Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA   Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA   Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA   Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA   Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA   Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA   Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA   Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA   Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA   Liu X., Mattei B., Mcintosh T.C., Mcleod M.P., Mcpherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA   Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H.,
RA   Wang X., Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA   Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT   "The genome sequence of Drosophila melanogaster.";
RL   Science 287:2185-2195(2000).
DR   EMBL; AE003601; AAF57770.1; -.
DR   FlyBase; FBgn0034286; CG12567.
DR   InterPro; IPR003598; Ig_c2.
DR   InterPro; IPR003006; Ig_MHC.
DR   Pfam; PF00047; ig; 2.
DR   SMART; SM00408; IGc2; 1.
KW   Immunoglobulin domain.
SQ   SEQUENCE   140 AA;  16191 MW;  05216CD136674A0 CRC64;
```

```
  Query Match             72.1%;  Score 31;  DB 5;  Length 140;
  Best Local Similarity   66.7%;  Pred. No. 43;
  Matches    6;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 FSKIASNIQ 9
           |: :|||
Db     118 FEVVASNIQ 126


RESULT 15
```

```
Q92DE9
ID   Q92DE9      PRELIMINARY;      PRT;   279 AA.
AC   Q92DE9;
DT   01-DEC-2001 (TrEMBLrel. 16, Created)
DT   01-DEC-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 16, Last annotation update)
DE   LIN0864 PROTEIN
GN   LIN0864.
OS   Listeria innocua.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group;
OC   Bacillus/Staphylococcus group; Listeria.
OX   NCBI_TaxID=1642;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN CLIP 11262 / SER:VAR 6A;
RX   PubMed=11679669;
RA   Glaser P., Frangeul L., Buchrieser C., Rusniok C., Amend A.,
RA   Baquero F., Berche P., Bloecker H., Brandt P., Chakraborty T.,
RA   Charbit A., Chetouani F., Couve E., de Daruvar A., Dehoux P.,
RA   Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,
RA   Entian K.-D., Fsihi H., Garcia-del Portillo F., Garrido P.,
RA   Gautier L., Goebel W., Gomez Lopez N., Hain T., Hauf J., Jackson D.,
RA   Jones L.-M., Kaerst U., Kreft J., Kuhn M., Kunst F., Kurapkat G.,
RA   Madueno E., Maitournam A., Mata Vicente J., Ng E., Nedjari H.,
RA   Nordsiek G., Novella S., de Pablos B., Perez-Diaz J.-C.,
RA   Purcell R., Remmel B., Rose M., Schlueter T., Simoes N., Tierrez A.,
RA   Vazquez-Boland J.-A., Voss H., Wehland J., Cossart P.;
RT   "Comparative genomics of Listeria species.";
RL   Science 294:849-852(2001).
DR   EMBL; AL596166; CAC96096.1; -.
DR   ListiList; LIN0864; -.
KW   Complete proteome.
SQ   SEQUENCE   279 AA;  32120 MW;  BFC11C1EA4CDE84 CRC64;
```

```
  Query Match             72.1%;  Score 31;  DB 5;  Length 279;
  Best Local Similarity   66.7%;  Pred. No. 89;
  Matches    6;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 FSKIASNIQ 9
           |: :|||
Db      80 FEKLVSNIQ 88


Search completed: August 26, 2002, 00:28:08
Job time: 1034 sec
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

M protein - protein search, using sw model

Run on: August 26, 2002, 06:13:24 ; Search time 139.28 Seconds
(without alignments)
7.177 Million cell updates/sec

Title: US-09-744-282-4
Perfect score: 43
Sequence: 1 FSKIASNIQ 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747971 seqs, 111135796 residues

Total number of hits satisfying chosen parameters: 747971

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_GenSeq_032802:*

1:	/cgn2_6/ptodata/1/pubaa/	1	protein	g/protein	AA.*
2:	/cgn2_6/ptodata/1/pubaa/	11	protein	g/protein	AA1981.DAT:*
3:	/cgn2_6/ptodata/1/pubaa/	3	protein	g/protein	AA1982.DAT:*
4:	/cgn2_6/ptodata/1/pubaa/	13	protein	g/protein	AA1983.DAT:*
5:	/cgn2_6/ptodata/1/pubaa/	13	protein	g/protein	AA1984.DAT:*
6:	/cgn2_6/ptodata/1/pubaa/	1	protein	g/protein	AA1985.DAT:*
7:	/cgn2_6/ptodata/1/pubaa/	13	protein	g/protein	AA1986.DAT:*
8:	/cgn2_6/ptodata/1/pubaa/	13	protein	g/protein	AA1987.DAT:*
9:	/cgn2_6/ptodata/1/pubaa/	13	protein	g/protein	AA1988.DAT:*
10:	/cgn2_6/ptodata/1/pubaa/	14	protein	g/protein	AA1989.DAT:*
11:	/cgn2_6/ptodata/1/pubaa/	14	protein	g/protein	AA1990.DAT:*
12:	/cgn2_6/ptodata/1/pubaa/	14	protein	g/protein	AA1991.DAT:*
13:	/cgn2_6/ptodata/1/pubaa/	14	protein	g/protein	AA1992.DAT:*
14:	/cgn2_6/ptodata/1/pubaa/	14	protein	g/protein	AA1993.DAT:*
15:	/cgn2_6/ptodata/1/pubaa/	14	protein	g/protein	AA1994.DAT:*
16:	/cgn2_6/ptodata/1/pubaa/	14	protein	g/protein	AA1995.DAT:*
17:	/cgn2_6/ptodata/1/pubaa/	14	protein	g/protein	AA1996.DAT:*
18:	/cgn2_6/ptodata/1/pubaa/	14	protein	g/protein	AA1997.DAT:*
19:	/cgn2_6/ptodata/1/pubaa/	14	protein	g/protein	AA1998.DAT:*
20:	/cgn2_6/ptodata/1/pubaa/	14	protein	g/protein	AA1999.DAT:*
21:	/cgn2_6/ptodata/1/pubaa/	14	protein	g/protein	AA2000.DAT:*
22:	/cgn2_6/ptodata/1/pubaa/	14	protein	g/protein	AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	43	100.0	9	17	AAR65274	Human cartilage gl
2	43	100.0	9	21	AAY64229	Peptide fragment o
3	43	100.0	13	17	AAR65078	Peptide conta. bf
4	43	100.0	13	21	AAY64243	Peptide fragment o
5	43	100.0	19	18	AAW47113	Human cartilage gl
6	43	100.0	19	18	AAW47114	Human cartilage gl
7	43	100.0	163	22	AAU02074	Synthetic human ta
8	43	100.0	179	21	AAG00227	Human secreted pro
9	43	100.0	210	22	AAU02078	Synthetic human mu
10	43	100.0	287	22	AAU02079	Synthetic multi ta
11	43	100.0	411	9	AAI81842	Polypeptide involv
12	43	100.0	450	22	AAU02076	Synthetic human B.
13	43	100.0	483	21	AAB03442	Gp38K protein sequ
14	43	100.0	459	22	AAU02077	Synthetic multi ta
15	33	76.7	383	18	AAW26751	Bovine whey pr 2nd
16	33	76.7	576	22	AAM80174	Human protein SEQ
17	33	76.7	576	22	AAM80175	Human protein SEQ
18	33	76.7	576	22	AAM41932	Human polypeptide
19	33	76.7	603	22	AAM79191	Human protein SEQ
20	33	76.7	636	22	AAM40146	Human polypeptide
21	33	76.7	650	22	AAM79198	Human protein SEQ
22	32	74.4	151	22	AAO02447	Human polypeptide
23	31	72.1	116	22	ABG04866	Novel human diagno
24	31	72.1	140	22	ABB69448	Drosophila melanog
25	31	72.1	179	22	AAM81813	Human haematopoiet
26	31	72.1	187	22	AAM82003	Human haematopoiet
27	31	72.1	292	22	ABG04870	Novel human diagno
28	31	72.1	406	22	AAE10163	RNA helicase cdna5
29	31	72.1	1216	22	ABB60349	Drosophila melanog
30	30	69.8	94	22	ABB10733	Human pancreatic p
31	30	69.8	94	22	AAM93031	Human diabetes sy
32	30	69.8	117	22	AAM40770	Human polypeptide
33	30	69.8	153	22	ABG07997	Novel human diagno
34	30	69.8	323	21	AAG53963	Arabidopsis thalia
35	30	69.8	432	21	AAG40048	Arabidopsis thalia
36	30	69.8	432	21	AAG48692	Arabidopsis thalia
37	30	69.8	445	21	AAG40047	Arabidopsis thalia
38	30	69.8	485	21	AAG30001	Arabidopsis thalia
39	30	69.8	485	21	AAG51798	Arabidopsis thalia
40	30	69.8	524	21	AAG30000	Arabidopsis thalia
41	30	69.8	524	21	AAG51797	Arabidopsis thalia
42	30	69.8	528	21	AAG22999	Arabidopsis thalia
43	30	69.8	531	21	AAG51796	Arabidopsis thalia
44	30	69.8	760	22	ABG14151	Novel human diagno
45	29	67.4	85	21	AAB27784	Protein fragment e

ALIGNMENTS

RESULT 1
AAR65274
ID AAR65274 standard; peptide; 9 AA.
XX
AC AAR65274;
XX
DT 17-DEC-1996 (first entry)
XX
DE Human cartilage gp-39 protein 9-mer antigenic determinant peptide.
XX
KW Human, cartilage, glycoprotein 39, gp-39; autoantigen;
KW Induction, T cell, tolerance, mediated, destruction;
KW rheumatoid arthritis, disease model, diagnosis.
XX
OS Homo sapiens.
XX
PN WO9613517-A1.
XX
PD 09-MAY-1996.
XX
PF 25-OCT-1995; 95WO-EP04201.
XX
PR 07-APR-1995; 95EP-0200886.
PR 27-OCT-1994; 94EP-0203128.
XX
PA (AKZO) AKZO NOBEL NV.
XX
PI Boots AMH, Verheijden GFM;
XX
DR WPI; 1996-239447/24.
XX
PT Peptide(s) derived from human cartilage gp-39 antigen used
PT to induce specific T cell tolerance, esp. for treatment of
PT rheumatoid arthritis, also useful in producing arthritic animal.

PT models
XX
PS Claim 1; Page 28; 87pp; English.
XX
CC The present peptide, derived from the human cartilage glycoprotein
CC 39 (hC gp 39) autoantigen, can be used to induce specific T cell
CC tolerance to hC gp 39 in a patient with T cell mediated cartilage
CC destruction, esp. rheumatoid arthritis (RA). It may also be used to
CC induce RA in mice, etc. (human disease models), and to diagnose the
CC presence of activated T cells autoreactive with hC gp 39, i.e. to
CC identify patients suitable for tolerising treatment.
CC Therapeutically the peptide is given in a dose of 0.1 100 microg/kg
CC by injection, or 10 50 microg/kg to produce a disease model. The
CC peptide has a specific tolerising effect on autoreactive T cells,
CC but leaves other components of the immune system intact, and
CC therefore should have few side effects. In an assay to determine
CC the peptide's binding to HLA-DR mols., specifically DR4Dw4 (known
CC to be associated with an increased risk of RA development), the
CC peptide gave a 50% blot IC(50) (micromolar) value of 0.08 compared
CC to 0.5% for the control peptide HA(307-319)F.
XX
SQ Sequence 9 AA;

Query Match 100.0%; Score 43; DB 17; Length 9;
Best Local Similarity 100.0%; Pred. No. 6.4e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FSKIASNTQ 9
||||||||| |
Db 1 fskiasntq 9

RESULT 2
AAY69229
ID AAY69229 standard; peptide; 9 AA.
XX
AC AAY69229;
XX
DT 30-MAY 2000 (first entry)
XX
DE Peptide fragment of human cartilage gp-39 (HC gp-39) protein.
XX
KW Human cartilage gp 39 protein; HC gp-39 protein; T cell ist vator;
KW inflammatory disease; lymphocyte; antigen, bystander suppression;
KW antigen specific T cell tolerance; autoimmune disorder;
KW inflammatory disease; rheumatoid arthritis;
KW autoimmune inflammatory disease; autoantigen.
XX
OS Homo sapiens.
XX
PN WO200004917-A2.
XX
PD 03-FEB-2000.
XX
PF 19-JUL-1999; 99WO-EP05431.
XX
PR 23-JUL-1998; 98EP-0202471.
XX
PA (ALKU) AKZO NOBEL NV.
XX
PI Miltenburg AMM, Boots AMH;
XX
DR WPI; 2000-182547/16.
XX
PT Use of cartilage protein or peptides for treating inflammatory diseases
PT especially rheumatoid arthritis
XX
PS Claim 2; Page 12; 29pp; English.
XX
CC AAY69226-33 represent fragments of the human cartilage gp 39 (HC gp 39)
CC protein. The peptides are T cell activators. The peptides are used for
CC preventing inflammatory diseases, by modulating lymphocytes reactive
CC to antigens other than HC gp 39, present in the same tissue as
CC HC gp-39. By the induction of antigen specific T cell tolerance,
CC autoimmune disorders can be treated by bystander suppression. The
CC HC gp-39 peptides are useful for preventing inflammation, disease,
CC especially rheumatoid arthritis. The fragments may also be used for
CC treating autoimmune inflammatory diseases. The HC gp 39 peptides are
CC especially beneficial in autoimmune conditions in which the
CC autoantigens associated with the condition are unknown.
XX
SQ Sequence 9 AA;

Query Match 100.0%; Score 43; DB 21; Length 9;
Best Local Similarity 100.0%; Pred. No. 6.4e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FSKIASNTQ 9
||||||||| |
Db 1 fskiasntq 9

RESULT 3
AAR95078
ID AAR95078 standard; peptide; 14 AA.
XX
AC AAR95078;
XX
DT 17-DEC-1996 (first entry)
XX
DE Peptide contg. hC gp 39 autoantigen derived peptide (IV).
XX
KW Human; cartilage; glycoprotein 39; hC gp 39; autoantigen;
KW induction; T-cell; tolerance; mediated; destruction;
KW rheumatoid arthritis; disease model; diagnosis.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Peptide 4..12
FT /note= "HC gp 39 peptide (IV)"
XX
PN WO9613517-A1.
XX
PD 09-MAY-1996.
XX
PF 25-OCT-1995; 95WO-EP04201.
XX
PR 27-OCT-1994; 94EP-0203128.
XX
PA (ALKU) AKZO NOBEL NV.
XX
PI Boots AMH, Verheijden GFM;
XX
DR WPI; 1996-239447/24.
XX
PT Peptide(s) derived from human cartilage gp 39 autoantigen useful
PT to induce specific T cell tolerance, esp. for treatment of
PT rheumatoid arthritis, also useful in producing arthritic animal
PT models
XX
PS Claim 2; Page 28; 87pp; English.
XX
CC The present peptide, contg. a peptide derived from the human
CC cartilage glycoprotein 39 (hC gp 39) autoantigen, can be used to
CC induce specific T cell tolerance to hC gp 39 in a patient with
CC T cell mediated cartilage destruction, esp. rheumatoid arthritis
CC (RA). It may also be used to induce RA in mice, etc. (human disease
CC models), and to diagnose the presence of activated T cells
CC autoreactive with hC gp 39, i.e. to identify patients suitable for
CC tolerising treatment.
CC Therapeutically the peptide is given in a dose of 0.1 100 microg/kg
CC by injection, or 10 50 microg/kg to produce a disease model. The

```
CC   peptide has a specific tolerising effect on autoreactive T-cells,
CC   but leaves other components of the immune system intact, and
CC   therefore should have few side effects.
XX
SQ   Sequence   14 AA


  Query Match             100.0%;  Score 43;  DB 17;  Length 14;
  Best Local Similarity   100.0%;  Pred. No. 0.01;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 FSKIASNIQ 9
           |||||||||
Db       4 tskiasntq 12


RESULT   4
AAY69233
ID   AAY69233 standard; peptide; 14 AA.
XX
AC   AAY69233;
XX
DT   30-MAY 2000  (first entry)
XX
DE   Peptide fragment of human cartilage gp-39 (HC gp-39) protein.
XX
KW   Human cartilage gp 39 protein; HC gp-39 protein; T-cell activator;
KW   inflammatory disease; lymphocyte; antigen; bystander suppression;
KW   antigen specific T-cell tolerance; autoimmune disorder;
KW   inflammatory disease; rheumatoid arthritis;
KW   autoimmune inflammatory disease; autoantigen.
XX
OS   Homo sapiens.
XX
PN   WO200004912-A2.
XX
PD   03-FEB-2000.
XX
PF   19-JUL-1999;   99WO-EP05333.
XX
PR   23-JUL-1998;   98EP-0201447.
XX
PA   (ALKO ) AKZO NOBEL NV.
XX
PI   Miltenburg AMM;  Boots AMH;
XX
DR   WPI; 2000-161247/16.
XX
PT   Use of cartilage protein or peptides for treating inflammatory diseases
PT   especially rheumatoide arthritis
XX
PS   Claim 2; Page 17; 29pp; English.
XX
CC   AAY69226-35 represent fragments of the human cartilage gp-39 (HC gp-39)
CC   protein. The peptides are T-cell activators. The peptides are used for
CC   prevention inflammatory diseases, by modulating lymphocytes reactive
CC   to antigens other that HC gp-39, present in the same tissue of
CC   HC gp-39. By the induction of antigen specific T-cell tolerance,
CC   autoimmune disorders can be treated by bystander suppression. The
CC   HC gp-39 peptides are useful for preventing inflammatory diseases
CC   especially rheumatoid arthritis. The fragments may also be used for
CC   treating autoimmune inflammatory diseases. The HC gp-39 peptides are
CC   especially beneficial in autoimmune conditions in which the
CC   autoantigens associated with the condition are unknown.
XX
SQ   Sequence   14 AA;


  Query Match             100.0%;  Score 43;  DB 21;  Length 14;
  Best Local Similarity   100.0%;  Pred. No. 0.01;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 FSKIASNIQ 9
           |||||||||
Db       4 tskiasntq 12


RESULT   5
AAW45113
ID   AAW45113 standard; peptide; 16 AA.
XX
AC   AAW45113;
XX
DT   28-APR-1998  (first entry)
XX
DE   Human cartilage glycoprotein-39 derived peptide #14.
XX
KW   Articular cartilage; immunosuppressive therapy; antigen; autoantigen;
KW   immunological tolerance; T-cell; human cartilage glycoprotein-39;
KW   HC gp-39; rheumatoid arthritis; epitope.
XX
OS   Synthetic.
OS   Homo sapiens.
XX
PN   WO9740068-A1.
XX
PD   30-OCT-1997.
XX
PF   22-APR-1997;   97WO-EP02051.
XX
PR   24-APR-1996;   96EP-0201106.
XX
PA   (ALKO ) AKZO NOBEL NV.
XX
PI   Boots AMH;  Verheijden GFM;
XX
DR   WPI; 1997-535775/49.
XX
PT   Peptide suitable for use in antigen specific immunosuppressive
PT   therapy - resembles or mimics epitope present on HC gp-39, so
PT   inducing systemic immunological tolerance to rheumatoid arthritis
PT   autoantigen
XX
PS   Disclosure; Page 15; 82pp; English.
XX
CC   The present sequence represents a peptide which resembles or mimics an
CC   epitope present on human cartilage glycoprotein 39 (HC gp 39), an
CC   autoantigen in rheumatoid arthritis. The invention relates to peptides
CC   consisting of 16-55 amino acid residues comprising at least one of the
CC   following 19 sequences: LVCYYISWS; FLCIHIIYS; IIYSFANIS; IKIIISVGG;
CC   FIKSVFFEL; FHGLILAWL; LYEGRRQSQ; YDIAKISQH; IDFISIMIY; FISIMIYIF;
CC   FRGQEDASP; YAVGYMLRL; MLRIGAPAS; LAYYFICDF; LHGAIVHRI; YIKHFLIAI;
CC   LACAMVWAL; VWALILDDF; or LDLIHPDGS. They can be used medically in
CC   antigen specific immunosuppressive therapy, particularly the treatment
CC   of T-cell mediated destruction of articular cartilage in autoimmune
CC   diseases (e.g. rheumatoid arthritis). They can also be used to detect
CC   activated autoreactive T cells in an individual. The peptides have a
CC   specific effect on the autoreactive T cells, thus leaving the other
CC   components of the immune system intact, unlike the non-specific
CC   suppressive effect of immunosuppressive drugs, and do not cause toxic
CC   side effects. The peptides are predominantly recognised by autoreactive
CC   T cells from rheumatoid arthritis patients, but rarely by those from
CC   healthy donors.
XX
SQ   Sequence   16 AA;


  Query Match             100.0%;  Score 43;  DB 18;  Length 16;
  Best Local Similarity   100.0%;  Pred. No. 0.012;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 FSKIASNIQ 9
           |||||||||
Db       7 tskiasntq 15
```

```
RESULT   6
AAW45114
ID   AAW45114 standard; peptide; 26 AA.
XX
AC   AAW45114;
XX
DT   28-APR-1998  (first entry)
XX
DE   Human cartilage glycoprotein 39 derived peptide #15.
XX
KW   Articular cartilage; immunosuppressive therapy; antigen; autoantigen;
KW   immunological tolerance; T cell; human cartilage glycoprotein 39;
KW   HC gp 39; rheumatoid arthritis; epitope.
XX
OS   Synthetic.
OS   Homo sapiens.
XX
PN   WO9740068-A1.
XX
PD   30-OCT-1997.
XX
PF   22-APR-1997;   97WO-EP02051.
XX
PR   24-APR-1996;   96EP-0201106.
XX
PA   (AKZO ) AKZO NOBEL NV.
XX
PI   Boots AMH;  Verheijden GFM;
XX
DR   WPI; 1997-545775/49.
XX
PT   Peptide suitable for use in antigen specific immunosuppressive
PT   therapy - resembles or mimics epitope present on HC gp 39, and
PT   induces systemic immunological tolerance to rheumatoid arthritis
PT   auto-antigen
XX
PS   Disclosure; Page 15; 82pp; English.
XX
CC   The present sequence represents a peptide which resembles or mimics an
CC   epitope present on human cartilage glycoprotein 39 (HC gp 39), an
CC   autoantigen in rheumatoid arthritis. The invention relates to peptides
CC   consisting of 16-55 amino acid residues comprising at least one of the
CC   following 16 sequences: LVCYYTSWS; FLCTHIIYS; IIYSFANIS; LKILLSVGG;
CC   FIFSVIFFI; FISSGGAW; CYCPPPPPG; YDIAKISIR; LDHSIMIY; FIFIMTYDF;
CC   FRGQEDASP; YAVGYMLFL; MLFIGAPAS; LAYYEICDF; LRGAIVHFI; YLKIEGLAI;
CC   LAGAMVWAI; VWALDLDDF; or LDLDDFQGS. They can be used medically in
CC   antigen specific immunosuppressive therapy, particularly the treatment
CC   of T cell mediated destruction of articular cartilage in autoimmune
CC   diseases (e.g. rheumatoid arthritis). They can also be used to detect
CC   activated autoreactive T cells in an individual. The peptides have a
CC   specific effect on the autoreactive T cells, thus leaving the other
CC   components of the immune system intact, unlike the non-specific
CC   suppressive effect of immunosuppressive drugs, and do not cause toxic
CC   side effects. The peptides are predominantly recognised by autoreactive
CC   T cells from rheumatoid arthritis patients, but rarely by those from
CC   healthy donors.
XX
SQ   Sequence   16 AA;


  Query Match             100.0%;  Score 43;  DB 18;  Length 16;
  Best Local Similarity   100.0%;  Pred. No. 0.012;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 FSKIASNTQ 9
           |||||||||
Db       1 fskiasntq 9


RESULT   7
AAG02074
ID   AAG02074 standard; Protein; 105 AA.
XX
AC   AAG02074;
XX
DT   07-SEP-2001  (first entry)
XX
DE   Synthetic human target autoantigen SHGLP/RA.
XX
KW   shGLP/RA; immunogenic epitope cluster; IEC;
KW   synthetic human chondrocyte glycoprotein 69/cartilage link protein;
KW   autoantigen; autoimmune disease; multiple sclerosis;
KW   rheumatoid arthritis; insulin dependent diabetes mellitus;
KW   myasthenia gravis; uveitis; autoimmune hepatitis; thyroiditis;
KW   orchitis; idiopathic thrombocytopenic purpura; inflammatory bowel disease;
KW   Crohn's disease; ulcerative colitis.
XX
OS   Synthetic.
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 103
FT                   /label= OTHER
FT                   /note= "Other: STOP"
XX
PN   WO200131037-A2
XX
PD   03-MAY-2001.
XX
PF   26-OCT-2000; 2000WO-IL00688.
XX
PR   27-OCT-1999;   99IL-0132611.
XX
PA   (YEDA ) YEDA RES & DEV CO LTD.
XX
PI   Ben Nun A;  Kerlero De Rosbo N;  Sappier GF;
XX
DR   WPI; 2001-300515/31.
DR   N-PSDB; AAS04827.
XX
PT   Novel synthetic human target autoantigen gene useful for treating
PT   autoimmune diseases such as multiple sclerosis, insulin dependent
PT   diabetes mellitus, rheumatoid arthritis, myasthenia gravis, uveitis
PT   -
XX
PS   Claim 173; Fig 51; 182pp; English.
XX
CC   The sequence represents shGLP/RA (synthetic human chondrocyte
CC   glycoprotein 69/cartilage link protein) containing immunogenic epitope
CC   clusters (IEC) from chondrocyte glycoprotein 69 and cartilage link
CC   protein. The synthetic human target autoantigen genes of the invention
CC   comprise sequences coding for at least 2 IECs of autoantigen(s) related
CC   to a specific autoimmune disease. The synthetic human target autoantigen
CC   genes are useful for treating autoimmune diseases such as multiple
CC   sclerosis, insulin dependent diabetes mellitus, rheumatoid arthritis,
CC   myasthenia gravis, uveitis, autoimmune hepatitis, thyroiditis, orchitis,
CC   orchitis, idiopathic thrombocytopenic purpura, and inflammatory bowel disease
CC   (Crohn's disease, ulcerative colitis). The synthetic human target
CC   autoantigen genes are also useful for diagnosis and/or monitoring the
CC   progression of the autoimmune disease.
XX
SQ   Sequence   105 AA;


  Query Match             100.0%;  Score 43;  DB 22;  Length 105;
  Best Local Similarity   100.0%;  Pred. No. 0.1;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 FSKIASNTQ 9
           |||||||||
Db      11 fskiasntq 19


RESULT   8
AAG30227
ID   AAG30227 standard; Protein; 126 AA.
```

```
XX
AC   AA002077;
XX
DT   06-OCT-2000  (first entry)
XX
DE   Human secreted protein, SEQ ID NO: 4308.
XX
KW   Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
KW   gene therapy; chromosome mapping.
XX
OS   Homo sapiens.
XX
PN   EP1033401-A2.
XX
PD   06-SEP-2000.
XX
PF   21-FEB-2000; 2000EP-0301610.
XX
PR   26-FEB-1999;  99US-0122487.
XX
PA   (GESE ) GENSET.
XX
PI   Dumas Milne Edwards J, Duclert A, Giordano J;
XX
DR   WPI; 2000-506380/45.
DR   N-PSDB; AAC00233.
XX
PT   New nucleic acid that is a 5' expressed sequence tag (5' EST) for
PT   obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
PT   diagnostic, forensic, gene therapy and chromosome mapping procedures
XX
PS   Claim 13; SEQ ID 4308; 71pp + CD-ROM; English.
XX
CC   The present sequence is a polypeptide encoded by one of a large number
CC   of 5' ESTs derived from mRNAs encoding secreted proteins. The 5' ESTs
CC   were prepared from total human RNAs or polyA+ RNAs derived from 30
CC   different tissues. EST sequences usually correspond mainly to the 3'
CC   untranslated region (UTR) of the mRNA because they are often obtained
CC   from oligodT primed cDNA libraries. Such ESTs are not well suited for
CC   isolating cDNA sequences derived from the 5' ends of mRNAs and even in
CC   those cases where 5 more cDNA sequences have been obtained, the full 5'
CC   UTR is rarely isolated. 5' ESTs are derived from mRNAs with intact 5'
CC   ends and can therefore be used to obtain full length cDNAs and genomic
CC   DNAs. 5' ESTs are also used in diagnostic, forensic, gene therapy and
CC   chromosome mapping procedures. They are used to obtain upstream
CC   regulatory sequences and to design expression and secretion vectors
XX
SQ   Sequence   126 AA;


  Query Match             100.0%;  Score 43;  DB 21;  Length 126;
  Best Local Similarity   100.0%;  Pred. No. 0.12;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 FSKIASNIQ 9
             |||||||||
Db     106 fskiasnig 114


RESULT   9
AA002078
ID   AA002078 standard; Protein; 220 AA.
XX
AC   AA002078;
XX
DT   07-SEP-2001  (first entry)
XX
DE   Synthetic human multi target autoantigen Y-RAPc.
XX
KW   Y-RAPc; immunogenic epitope cluster; IEC;
KW   autoimmune hepatitis; thyroiditis; insulitis; orchitis;
KW   idiopathic thrombocytopenic purpura; inflammatory disease;
KW   Crohn's disease; ulcerative colitis; shMultiTAG.
XX
OS   Synthetic.
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Misc difference 218
FT                   /label= OTHER
FT                   /note= "Other- SIEP"
XX
PN   WO200141037-A2.
XX
PD   03-MAY-2001.
XX
PF   26-OCT-2000; 2000WO-IL000688.
XX
PR   27-OCT-1999;  99IL-0132611.
XX
PA   (YEDA ) YEDA RES & DEV CO LTD.
XX
PI   Ben Nun A, Kerlero De Rosbo N, Sappler GP;
XX
DR   WPI; 2001-300515/31.
DR   N-PSDB; AAS04839.
XX
PT   Novel synthetic human target autoantigen gene useful for treating
PT   autoimmune diseases such as multiple sclerosis, insulin dependent
PT   diabetes mellitus, rheumatoid arthritis, myasthenia gravis, and uveitis
PT   -
XX
PS   Claim 185; Fig 59; 182pp; English.
XX
CC   The sequence represents synthetic human multi-target autoantigen
CC   (shMultiTAG) Y-RAPa consisting of regions encoding preferred epitopes
CC   from shGelsolin/hA, shAquaporin/hA and shGFP/hA. The synthetic human
CC   target autoantigen genes of the invention comprise sequences coding for
CC   at least 2 IECs of autoantigen(s) related to a specific autoimmune
CC   disease. The synthetic human target autoantigen genes are useful for
CC   treating autoimmune diseases such as multiple sclerosis, insulin-
CC   dependent diabetes mellitus, rheumatoid arthritis, myasthenia gravis,
CC   uveitis, autoimmune hepatitis, thyroiditis, insulitis, orchitis,
CC   idiopathic thrombocytopenic purpura, and inflammatory diseases (Crohn's
CC   disease, ulcerative colitis). The synthetic human target autoantigen
CC   genes are also useful for diagnosis and/or monitoring the progression of
CC   the autoimmune disease.
XX
SQ   Sequence   220 AA;


  Query Match             100.0%;  Score 43;  DB 22;  Length 220;
  Best Local Similarity   100.0%;  Pred. No. 0.24;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 FSKIASNIQ 9
             |||||||||
Db     167 fskiasnig 175


RESULT   10
AA002079
ID   AA002079 standard; Protein; 281 AA.
XX
AC   AA002079;
XX
DT   07-SEP-2001  (first entry)
XX
DE   Synthetic multi target autoantigen Y-RAPd.
XX
KW   Y-RAPd; immunogenic epitope cluster; IEC;
KW   autoimmune hepatitis; thyroiditis; insulitis; orchitis;
KW   autoimmune disease; multiple sclerosis; rheumatoid arthritis;
```

KW insulin-dependent diabetes mellitus; myasthenia gravis; uveitis;
KW autoimmune hepatitis; thyroiditis; insulitis; orchitis;
KW idiopathic thrombocytopenic purpura; inflammatory disease;
KW Crohn's disease; ulcerative colitis; shMultiTAG.
XX
OS Synthetic.
OS Homo sapiens.
OS Escherichia coli.
OS Mycobacterium tuberculosis.
XX
FH Key Location/Qualifiers
FT Misc difference 279
FT /label= OTHER
FT /note= "Other= STOP"
XX
PN WO200131037 A2.
XX
PD 03 MAY 2001.
XX
PF 26 OCT 2000; 2000WO-IL00688.
XX
PR 27 OCT 1999; 99IL-0132611.
XX
PA (YEDA) YEDA RES & DEV CO LTD.
XX
PI Ben Nun A, Kerlero De Rosbo N, Sappier GP;
XX
DR WPI; 2001-305115/31.
DR N-PSDB; AAS04840.
XX
PT Novel synthetic human target autoantigen gene useful in treating
PT autoimmune diseases such as multiple sclerosis, insulin dependent
PT diabetes mellitus, rheumatoid arthritis, myasthenia gravis, and uveitis
PT
XX
PS Claim 108; Fig 61; 182pp; English.
XX
CC The sequence represents synthetic human multi target autoantigen
CC (shMultiTAG) Y kAPd consisting of preferred epitopes encoded by
CC shCollagen XI, shAlpha-Cryst A, shMyelin PLP, shMyelin MBP, Insu...
CC target autoantigen genes of the invention comprise sequences coding for
CC at least 2 IE's of autoantigen(s) related to a specific autoimmune
CC disease. The synthetic human target autoantigen genes are useful for
CC treating autoimmune diseases such as multiple sclerosis, insulin
CC dependent diabetes mellitus, rheumatoid arthritis, myasthenia gravis,
CC uveitis, autoimmune hepatitis, thyroiditis, insulitis, orchitis,
CC idiopathic thrombocytopenic purpura, and inflammatory disease (Crohn's
CC disease, ulcerative colitis). The synthetic human target autoantigen
CC genes are also useful for diagnosis and/or monitoring the progression of
CC the autoimmune disease.
XX
SQ Sequence 281 AA;

Query Match 100.0%; Score 43; DB 22; Length 281;
Best Local Similarity 100.0%; Pred. No. 0.41;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FSKIASNIQ 9
||||||||| |
Db 167 fskiasnlq 175

RESULT 11
AAP81442
ID AAP81442 standard; protein; 321 AA.
XX
AC AAP81442;
XX
DT 19-OCT-1990 (first entry)
XX
DE Polypeptide involved in protective mechanisms.
XX
KW Immune response; cell growth.
XX
PN JP63042898-A.
XX
PD 30-JAN-1988.
XX
PF 26-JUL-1986; 86JP-0175518.
XX
PR 16-JUL-1986; 86JP-0167518.
XX
PA (JAPA-) JAPAN IBARAKI KK.
XX
DR WPI; 1988-068419/10.
DR P-PSDB; AAP81442.
XX
PT New polypeptide and DNA encoding it - related to protective
PT mechanisms such as immune response etc.
XX
PS Disclosure; ; pp; Japanese.
XX
CC This polypeptide is involved in protective mechanisms such as immune
CC response, cell growth and activation of protective functions.
XX
SQ Sequence 321 AA;

Query Match 100.0%; Score 43; DB 2; Length 321;
Best Local Similarity 100.0%; Pred. No. 0.46;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FSKIASNIQ 9
||||||||| |
Db 165 fskiasnlq 173

RESULT 12
AAU02076
ID AAU02076 standard; Protein; 450 AA.
XX
AC AAU02076;
XX
DT 07-SEP-2001 (first entry)
XX
DE Synthetic human multi target autoantigen Y kAPd.
XX
KW Y kAPd; immunodominant epitope cluster; IE;
KW autoimmune disease; multiple sclerosis; rheumatoid arthritis;
KW insulin-dependent diabetes mellitus; myasthenia gravis; uveitis;
KW autoimmune hepatitis; thyroiditis; insulitis; orchitis;
KW idiopathic thrombocytopenic purpura; inflammatory disease;
KW Crohn's disease; ulcerative colitis; shMultiTAG.
XX
OS Synthetic.
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Misc-difference 448
FT /label= OTHER
FT /note= "Other= STOP"
XX
PN WO200131037 A2.
XX
PD 03 MAY 2001.
XX
PF 26 OCT-2000; 2000WO-IL00688.
XX
PR 27 OCT 1999; 99IL-0132611.
XX
PA (YEDA) YEDA RES & DEV CO LTD.
XX
PI Ben Nun A, Kerlero De Rosbo N, Sappier GP;
XX

```
DR   WPI; 2000-431300/37.
XX
PT   Clusterin and qp58K-related peptide capable of altering cell migration
PT   useful for treating atherosclerosis, cancer and stenosis following
PT   vascular trauma or disease
XX
PS   Disclosure; Fig 2; 43pp; English.
XX
CC   The present sequence is the protein sequence of qp58k. Qp58k, a
CC   chemoattractant, is essential for the migration of vascular smooth muscle
CC   cells (VSMC). The gene and protein can, therefore, be used to promote
CC   wound healing, angiogenesis and vasculogenesis, in the treatment of
CC   stenosis following vascular trauma or disease and to treat
CC   atherosclerosis, and antisense sequences can be used to treat cancer, as
CC   angiogenesis is vital for tumour survival.
XX
SQ   Sequence   383 AA;


  Query Match             100.0%;  Score 43;  DB 21;  Length 383;
  Best Local Similarity   100.0%;  Pred. No. 0.45;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 FSKIASNTQ 9
           |||||||||
Db     106 fskiasntq 114


RESULT  14
AAU02077
ID   AAU02077 standard; Protein; 459 AA.
XX
AC   AAU02077;
XX
DT   07-SEP-2001  (first entry)
XX
DE   Synthetic multi-target autoantigen Y-RAPb.
XX
KW   Y-RAPa; immunogenic epitope cluster; IEC;
KW   [illegible] shAspartoacylase; shMultiAG;
KW   autoimmune disease; multiple sclerosis; rheumatoid arthritis;
KW   insulin-dependent diabetes mellitus; myasthenia gravis; uveitis;
KW   autoimmune hepatitis; thyroiditis; insulitis; orchitis;
KW   idiopathic thrombocytopenic purpura; inflammatory disease;
KW   Crohn's disease; ulcerative colitis; shMultiAG.
XX
OS   Synthetic.
OS   Homo sapiens.
OS   Escherichia coli.
OS   Mycobacterium tuberculosis.
XX
FH   Key             Location/Qualifiers
FT   Misc difference 455
FT                   /label  OTHER
FT                   /note  "other  STOP"
FT   Misc-difference 457
FT                   /label  OTHER
FT                   /note  "other  STOP"
FT   Binding site    368..376
FT                   /label  hyaluronic_acid_binding_site
XX
PN   WO200084469 A1.
XX
PD   15-JUN-2000.
XX
PF   10-DEC-1999;  99WO-US29262.
XX
PR   11-DEC-1998;  98US-0111896.
XX
PA   (UYNY ) UNIV NEW YORK STATE RES FOUND.
XX
PI   Millis AJT;
XX
```

XX
PN WO200131037 A2.
XX
PD 03 MAY 2001.
XX
PF 26 OCT 2000; 2000WO-IL00688.
XX
PR 27 OCT 1999; 99IL-0132611.
XX
PA (YEDA) YEDA RES & DEV CO LTD.
XX
PI Ben-Nun A, Kerlero De Rosbo N, Sappler GP;
XX
DR WPI; 2001-300515/31.
DR N-PSDB; AAS04838.
XX
PT Novel synthetic human target autoantigen gene useful for treating
PT autoimmune diseases such as multiple sclerosis, insulin dependent
PT diabetes mellitus, rheumatoid arthritis, myasthenia gravis, and uveitis
PT
XX
PS Claim 182; Fig 57; 182pp; English.
XX
CC The sequence represents synthetic human multi-target autoantigen
CC (sHuMultiTAG) Y-PAPA consisting of regions encoding sHCollagen/EA,
CC sHAgrin-scEA, sHIg/EA and sHBP3/EA. The synthetic human
CC autoantigen genes of the invention comprise sequences coding for at least
CC 2 TE's of autoantigen(s) related to a specific autoimmune disease. The
CC synthetic human target autoantigen genes are useful for treating
CC autoimmune diseases such as multiple sclerosis, insulin-dependent
CC diabetes mellitus, rheumatoid arthritis, myasthenia gravis, uveitis,
CC autoimmune hepatitis, thyroiditis, insulitis, orchitis, idiopathic
CC thrombocytopenic purpura, and inflammatory diseases (Crohn's disease,
CC ulcerative colitis). The synthetic human target autoantigen genes are
CC also useful for diagnosis and/or monitoring the progression of the
CC autoimmune disease.
XX
SQ Sequence 459 AA;

Query Match 100.0%; Score 43; DB 22; Length 459;
Best Local Similarity 100.0%; Pred. No. 0.55;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FSKIASNTQ 9
||||||||| |
Db 256 fskiasntq 264

RESULT 15
AAW26751
ID AAW26751 standard; Protein; 383 AA.
XX
AC AAW26751;
XX
DT 11-MAY 1998 (first entry)
XX
DE Bovine whey protein.
XX
KW Bovine whey protein; human cartilage glycoprotein 39; HC gp-39;
KW autoantigen; antigen; autoimmune disease; rheumatoid arthritis;
KW inflammation; arthritogenic protein; immunotherapy; therapy.
XX
OS Bos taurus.
XX
PN WO9740149 A1.
XX
PD 30 OCT 1997.
XX
PF 15-APR-1997; 97WO-EP01903.
XX
PR 18-APR-1996; 96US-0634493.
XX
PA (ALKU) AKZO NOBEL NV.
XX
PI Boots AMH, Bos ES, Verheijden GFM;
XX
DR WPI; 1997-535833/49.
DR N-PSDB; AAT99452.
XX
PT Autoantigen proteins homologous to human cartilage glycoprotein 39
PT induce arthritis and provide antigen-specific treatment of arthritis or
PT cartilage destruction in autoimmune diseases e.g. rheumatoid
PT arthritis
XX
PS Claim 11; Page 23-25; 35pp; English.
XX
CC This bovine 39 kDa whey protein sequence was deduced from cDNA
CC clones (see AAT99452) isolated from a cattle mammary gland cDNA
CC library using human cartilage glycoprotein 39 (HC gp-39) cDNA as
CC probe. Bovine whey protein was shown to be arthritogenic, inducing
CC arthritis in animals in the same way as described for HC gp-39 (see
CC AAW26750). Arthritogenic proteins such as bovine whey protein, and
CC which display at least 50% amino acid homology to HC gp-39, are
CC very suitable for inducing systemic tolerance of the immune system
CC to homologous autoantigens and can be used to delay and/or suppress
CC arthritic development in mammals. They can induce specific long-
CC tolerance to HC gp-39 in patients suffering from T cell mediated
CC cartilage destruction, such as rheumatoid arthritis. Arthritogenic
CC proteins are also suitable to induce arthritis in animals,
CC preferably mice, e.g. for use in drug screening.
XX
SQ Sequence 383 AA;

Query Match 76.7%; Score 33; DB 19; Length 383;
Best Local Similarity 77.8%; Pred. No. 65;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FSKIASNTQ 9
|||||| :
Db 106 fsklasktr 114

Search completed: August 26, 2002, 06:14:25
Job time: 267 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

SW protein protein search, using sw model

Run on: August 26, 2002, 06:23:27 ; Search time 46.22 Seconds
(without alignments)
4,756 Million cell updates/sec

Title: US-09-744-282-4
Perfect score: 43
Sequence: 1 FSKIASNTQ 9

Scoring table: BLOSUM62
Gap 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iss_pep/*_SW.pep:*
2: /cgn2_6/ptodata/1/iss_pep/*_SW.pep:*
3: /cgn2_6/ptodata/1/iss_pep/*_SW.pep:*
4: /cgn2_6/ptodata/1/iss_pep/*_SW.pep:*
5: /cgn2_6/ptodata/1/iss_pep/*_SW.pep:*
6: /cgn2_6/ptodata/1/iss_pep/*_SW.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	43	100.0	9	1	US-08-619-645-4	Sequence 4, Appli
2	43	100.0	9	2	US-08-634-493-4	Sequence 4, Appli
3	43	100.0	14	1	US-08-129-645-8	Sequence 8, Appli
4	43	100.0	14	2	US-08-634-493-8	Sequence 8, Appli
5	43	100.0	16	4	US-09-171-705-14	Sequence 14, Appl
6	43	100.0	16	4	US-09-171-705-15	Sequence 15, Appl
7	29	67.4	176	4	US-08-478-413-34	Sequence 34, Appl
8	29	67.4	196	2	US-08-468-418-2	Sequence 2, Appli
9	28	65.1	144	4	US-08-630-915A-22	Sequence 22, Appl
10	28	65.1	151	1	US-08-435-454-4	Sequence 4, Appli
11	28	65.1	151	1	US-08-652-472A-4	Sequence 4, Appli
12	28	65.1	151	2	US-08-919-145-6	Sequence 6, Appli
13	28	65.1	151	3	US-08-870-126-4	Sequence 4, Appli
14	28	65.1	151	4	US-09-444-889-6	Sequence 6, Appli
15	28	65.1	151	5	PCT US96-10231A-4	Sequence 4, Appli
16	28	65.1	179	4	US-08-755-587-184	Sequence 184, App
17	28	65.1	2429	4	US-08-755-587-16	Sequence 16, Appl
18	28	65.1	3418	2	US-08-839-501-2	Sequence 2, Appli
19	28	65.1	3418	2	US-08-603-750-4	Sequence 4, Appli
20	28	65.1	3418	3	US-09-044-946-2	Sequence 2, Appli
21	28	65.1	3418	3	US-08-755-587-44	Sequence 44, Appl
22	28	65.1	3418	3	US-09-044-988-2	Sequence 2, Appli
23	28	65.1	3418	4	US-09-099-754-4	Sequence 4, Appli
24	28	65.1	3418	4	US-08-986-106-4	Sequence 4, Appli
25	27	62.8	109	4	US-09-240-274-54	Sequence 54, Appl
26	27	62.8	349	2	US-08-874-460-2	Sequence 2, Appli
27	27	62.8	385	3	US-08-633-993A-11	Sequence 11, Appl
28	27	62.8	385	3	US-08-844-188-11	Sequence 11, Appl
29	27	62.8	542	3	US-08-911-421-11	Sequence 11, Appl
30	27	62.8	542	5	PCT US95-13975-2	Sequence 2, Appl
31	27	62.8	981	2	US-08-649-046-2	Sequence 2, Appli
32	27	62.8	2285	4	US-09-308-375-2	Sequence 2, Appli
33	27	62.8	3169	2	US-08-477-451-6	Sequence 6, Appli
34	26	60.5	41	2	US-07-921-447-10	Sequence 10, Appl
35	26	60.5	157	3	US-08-851-843A-83	Sequence 83, Appl
36	26	60.5	157	4	US-08-974-549A-226	Sequence 226, App
37	26	60.5	157	4	US-08-854-050-83	Sequence 83, Appl
38	26	60.5	157	4	US-09-430-324-83	Sequence 83, Appl
39	26	60.5	233	3	US-08-851-843A-64	Sequence 64, Appl
40	26	60.5	233	4	US-08-974-549A-15	Sequence 15, Appl
41	26	60.5	233	4	US-08-854-050-64	Sequence 64, Appl
42	26	60.5	233	4	US-09-430-324-64	Sequence 64, Appl
43	26	60.5	253	4	US-08-944-483-73	Sequence 73, Appl
44	26	60.5	254	2	US-08-560-098A-49	Sequence 49, Appl
45	26	60.5	306	2	US-08-560-098A-45	Sequence 45, Appl

ALIGNMENTS

RESULT 1
US-08-619-645-4
; Sequence 4, Application US/08619645
; Patent No. 5736507
; GENERAL INFORMATION:
; APPLICANT: Boots, Anna M.H.
; APPLICANT: Verheijden, Gijsbertus F.M.
; TITLE OF INVENTION: [illegible] peptides derived from
; TITLE OF INVENTION: autoantigen for use in immunotherapy of autoimmune
; TITLE OF INVENTION: disease
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Akzo Nobel Patent Department
; STREET: 1400 Piccard Drive, Suite 206
; CITY: Rockville
; STATE: Maryland
; COUNTRY: U.S.A.
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25(EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/619,645
; FILING DATE: 25-MAR-1996
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Gormley, Mary E.
; REGISTRATION NUMBER: 34,409
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 947-4433
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-619-645-4

Query Match 100.0%; Score 43; DB 1; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.7e-05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FSKIASNTQ 9
|||||||||
Db 1 FSKIASNTQ 9

```
RESULT   2
US-08-634-493-4
; Sequence 4, Application US/08634493
; Patent No. 5844449
;  GENERAL INFORMATION:
;     APPLICANT:  A.M.H. Boots
;     APPLICANT:  G.F.M. Verheijden
;     APPLICANT:  E.S. Bos
;     TITLE OF INVENTION:  No. 5844449o1 Peptides derived from autoantigens for use
;     TITLE OF INVENTION:  in Immunotherapy of Autoimmune Diseases
;     NUMBER OF SEQUENCES:  10
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Akzo No. 5844449o1 Patent Department
;       STREET:  1300 Piccard Drive, Suite 206
;       CITY:  Rockville
;       STATE:  Maryland
;       COUNTRY:  USA
;       ZIP:  20850
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/634,493
;       FILING DATE:  18-APR-1996
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/619,645
;       FILING DATE:  25-MAR-1996
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  PCT/EP95/04201
;       FILING DATE:  25-OCT 1995
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  NL 94203287
;       FILING DATE:  27-OCT 1994
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  NL 95200860
;       FILING DATE:  07-APR 1995
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (301) 258-5200
;       TELEFAX:  (301) 977 0847
;   INFORMATION FOR SEQ ID NO:  4:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  9 amino acids
;       TYPE:  amino acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  peptide
US-08-634-493-4


  Query Match             100.0%;  Score 43;  DB 2;  Length 9;
  Best Local Similarity   100.0%;  Prod. No. 1.7e-05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 FSKIASNTQ 9
           |||||||||
Db       1 FSKIASNTQ 9


RESULT   3
US-08-619-645-8
; Sequence 8, Application US/08619645
; Patent No. 5746507
;  GENERAL INFORMATION:
;     APPLICANT:  Boots, Anna M.H.
;     APPLICANT:  Verheijden, Gijsbertus F.M.
;     TITLE OF INVENTION:  No. 5746507o1 peptides derived fr...
;     TITLE OF INVENTION:  autoantigens for use in immunotherap... autoimmune
;     TITLE OF INVENTION:  disease
;     NUMBER OF SEQUENCES:  9
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Akzo No. 5746507o1 Patent Department
;       STREET:  1300 Piccard Drive, Suite 206
;       CITY:  Rockville
;       STATE:  Maryland
;       COUNTRY:  U.S.A.
;       ZIP:  20850
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC DOS/MS DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25(EPO)
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/ 8/619,645
;       FILING DATE:  25-MAR 1996
;       CLASSIFICATION:  514
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Gormley, Mary E.
;       REGISTRATION NUMBER:  34,409
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (301) 947 4843
;   INFORMATION FOR SEQ ID NO:  8:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  14 amino acids
;       TYPE:  amino acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  peptide
US-08-619-645-8


  Query Match             100.0%;  Score 43;  DB 1;  Length 14;
  Best Local Similarity   100.0%;  Prod. No. 2.6e-07;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 FSKIASNTQ 9
           |||||||||
Db       4 FSKIASNTQ 12


RESULT   4
US-08-634-493-8
; Sequence 8, Application US/08634493
; Patent No. 5844449
;  GENERAL INFORMATION:
;     APPLICANT:  A.M.H. Boots
;     APPLICANT:  G.F.M. Verheijden
;     APPLICANT:  E.S. Bos
;     TITLE OF INVENTION:  No. 5844449o1 Peptides derived from autoantigens for use
;     TITLE OF INVENTION:  in Immunotherapy of Autoimmune Diseases
;     NUMBER OF SEQUENCES:  10
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Akzo No. 5844449o1 Patent Department
;       STREET:  1300 Piccard Drive, Suite 206
;       CITY:  Rockville
;       STATE:  Maryland
;       COUNTRY:  USA
;       ZIP:  20850
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/634,493
;       FILING DATE:  18-APR 1996
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/619,645
;       FILING DATE:  25-MAR 1996
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  PCT/EP95/04201
;       FILING DATE:  25-OCT 1995
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  NL 94203287
```

```
;       FILING DATE:  27-OCT-1994
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  NL 95200860
;       FILING DATE:  07-APR-1995
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (301) 258-5200
;       TELEFAX:  (301) 427-0947
;   INFORMATION FOR SEQ ID NO:  8:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  14 amino acids
;       TYPE:  amino acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  peptide
US-08-634-403-8


  Query Match             100.0%;  Score 43;  DB 2;  Length 14;
  Best Local Similarity   100.0%;  Pred. No. 0.0027;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 FSKIASNTQ 9
           |||||||||
Db       4 FSKIASNTQ 12


RESULT   5
US-09-171-705-14
; Sequence 14, Application US/09171705
; Patent No. 6184204
; GENERAL INFORMATION:
;  APPLICANT: BOOTS, ANNA M.H.
;  APPLICANT: VERHEIJDEN, GILBERTUS F.M.
;  TITLE OF INVENTION: NOVEL PEPTIDES SUITABLE FOR USE IN ANTIGEN SPECIFIC
;  TITLE OF INVENTION: IMMUNOSUPPRESSIVE THERAPY
;  FILE REFERENCE: O/96198 US
;  CURRENT APPLICATION NUMBER: US/09/171,705
;  CURRENT FILING DATE: 1999-02-09
;  NUMBER OF SEQ ID NOS: 78
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 14
;   LENGTH: 16
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: DERIVED FROM
;   OTHER INFORMATION: SEQUENCE OF HUMAN CARTILAGE (BC)  AT RESIDUE
US-09-171-705-14


  Query Match             100.0%;  Score 43;  DB 4;  Length 16;
  Best Local Similarity   100.0%;  Pred. No. 0.0041;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 FSKIASNTQ 9
           |||||||||
Db       7 FSKIASNTQ 15


RESULT   6
US-09-171-705-15
; Sequence 15, Application US/09171705
; Patent No. 6184204
; GENERAL INFORMATION:
;  APPLICANT: BOOTS, ANNA M.H.
;  APPLICANT: VERHEIJDEN, GILBERTUS F.M.
;  TITLE OF INVENTION: NOVEL PEPTIDES SUITABLE FOR USE IN ANTIGEN SPECIFIC
;  TITLE OF INVENTION: IMMUNOSUPPRESSIVE THERAPY
;  FILE REFERENCE: O/96198 US
;  CURRENT APPLICATION NUMBER: US/09/171,705
;  CURRENT FILING DATE: 1999-02-09
;  NUMBER OF SEQ ID NOS: 78
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 15
;   LENGTH: 16
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: DERIVED FROM
;   OTHER INFORMATION: SEQUENCE OF HUMAN CARTILAGE (BC)  AT RESIDUE
US-09-171-705-15


  Query Match             100.0%;  Score 43;  DB 4;  Length 16;
  Best Local Similarity   100.0%;  Pred. No. 0.0041;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 FSKIASNTQ 9
           |||||||||
Db       1 FSKIASNTQ 9


RESULT   7
US-08-378-313-34
; Sequence 34, Application US/08378313
; Patent No. 6207881
;  GENERAL INFORMATION:
;    APPLICANT:  THEOLOGIS, ATHANASIOS
;    APPLICANT:  SATO, TAKAHIDO
;    TITLE OF INVENTION:  CONTROL OF FRUIT RIPENING THROUGH
;    TITLE OF INVENTION:  GENETIC CONTROL OF ACC SYNTHASE SYNTHESIS
;    NUMBER OF SEQUENCES:  34
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  MORRISON & FOERSTER
;      STREET:  755 Page Mill Road
;      CITY:  Palo Alto
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94304-1018
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.2
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/378,313
;      FILING DATE:
;      CLASSIFICATION:  800
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/862,493
;      FILING DATE:  02-APR-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  MURASHIGE, KATE H.
;      REGISTRATION NUMBER:  29,959
;      REFERENCE/DOCKET NUMBER:  29190-200-2,20
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 856-5600
;      TELEFAX:  (415) 494-0792
;      TELEX:  706141
;  INFORMATION FOR SEQ ID NO:  34:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  476 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  281
;      OTHER INFORMATION:  /note= "This position is S or A."
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  284
;      OTHER INFORMATION:  /note= "This position is M or T."
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  286
```

```
        ; OTHER INFORMATION:  /note= "Xaa is Ser, Ile or Thr"
      FEATURE:
        NAME/KEY:  Modified-site
        LOCATION:  208
        ; OTHER INFORMATION:  /note= "Xaa is Ser, Ile or Thr"
US-08-744-282-4

  Query Match             67.4%;  Score 29;  DB 4;  Length 479;
  Best Local Similarity   75.0%;  Pred. No. 1.4e+02;
  Matches    6;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       4 SKIASNIQ 9
           ||| | |
Db     240 SKIASNLQ 247


RESULT  8
US-08-434-454-2
; Sequence 2, Application US/08463418
; Patent No. 5608971
;  GENERAL INFORMATION:
;    APPLICANT:  Van Der Straeten, Dominique et al.
;    TITLE OF INVENTION:  PROMOTER ACT SYNTHASE AND USES THEREOF
;    NUMBER OF SEQUENCES:  10
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fish & Richardson P.C.
;      STREET:  225 Franklin Street
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  USA
;      ZIP:  02110-2804
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/463,418
;      FILING DATE:  05-JUN-1995
;      CLASSIFICATION:  800
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/902,481
;      FILING DATE:  19-JUN-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Clark, Paul T.
;      REGISTRATION NUMBER:  30,162
;      REFERENCE/DOCKET NUMBER:  07470/204002
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617/542-5070
;      TELEFAX:  617/542-8906
;      TELEX:  200154
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  496 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-434-454-2


  Query Match             67.4%;  Score 29;  DB 4;  Length 496;
  Best Local Similarity   75.0%;  Pred. No. 1.4e+02;
  Matches    6;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       4 SKIASNIQ 9
           ||| | |
Db     240 SKIASNLQ 247


RESULT  9
US-08-630-915A-22
; Sequence 22, Application US/08630915A
; Patent No. 6309820
;  GENERAL INFORMATION:
;    APPLICANT:  SPARKS, Andrew B.
;    APPLICANT:  BOETMAN, Nic Gerardus
;    APPLICANT:  KAY, Brian K.
;    APPLICANT:  FOWLKES, Dana M.
;    APPLICANT:  McCONNELL, Stephen J.
;    TITLE OF INVENTION:  POLYPEPTIDES HAVING A FUNCTIONAL
;    TITLE OF INVENTION:  DOMAIN OF INTEREST AND METHODS OF IDENTIFYING AND
;    TITLE OF INVENTION:  USING SAME
;    NUMBER OF SEQUENCES:  227
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fennie & Edmonds LLP
;      STREET:  1155 Avenue of the Americas
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10036-2711
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/630,915A
;      FILING DATE:  03-APR-1996
;      CLASSIFICATION:  536
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Misrock, S. Leslie
;      REGISTRATION NUMBER:  18,872
;      REFERENCE/DOCKET NUMBER:  1101-174
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 790-9090
;      TELEFAX:  (212) 869-8864/9741
;      TELEX:  66141 FENNIE
;  INFORMATION FOR SEQ ID NO:  22:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  434 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:  peptide
US-08-630-915A-22


  Query Match             65.1%;  Score 28;  DB 4;  Length 434;
  Best Local Similarity   85.7%;  Pred. No. 1.9e+02;
  Matches    6;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       3 KIASNIQ 9
           ||| | |
Db      13 KIASNVQ 19


RESULT  10
US-08-435-454-4
; Sequence 4, Application US/08435454
  Patent No. 5605830
;  GENERAL INFORMATION:
;    APPLICANT:  Prendergast, George C.
;    APPLICANT:  Sakamuro, Daitoku
;    TITLE OF INVENTION:  Murine and Human C-Myc Interacting
;    TITLE OF INVENTION:  Protein and Uses Thereto
;    NUMBER OF SEQUENCES:  5
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Howson and Howson
;      STREET:  Spring House Corporate Cntr, P.O.Box 457
;      CITY:  Spring House
;      STATE:  Pennsylvania
;      COUNTRY:  USA
;      ZIP:  19477
;    COMPUTER READABLE FORM:
```

```
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.30
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/435,454
;       FILING DATE:
;       CLASSIFICATION:  514
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Bak, Mary E.
;       REGISTRATION NUMBER:  31,215
;       REFERENCE/DOCKET NUMBER:  WS1608USA
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  215-540-9200
;       TELEFAX:  215-540-5818
;   INFORMATION FOR SEQ ID NO:  4:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  451 amino acids
;       TYPE:  amino acid
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  protein
US-08-435-454-4


  Query Match             65.1%;  Score 28;  DB 1;  Length 451;
  Best Local Similarity   85.7%;  Pred. No. 2e+02;
  Matches    6;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       5 KIASNIQ 9
            ||||| |
Db      10 KIASNVQ 16


RESULT  11
US-08-652-972A-4
; Sequence 4, Application US/08652972A
; Patent No. 5723581
;  GENERAL INFORMATION:
;     APPLICANT:  Prendergast, George C.
;     APPLICANT:  Sakamuro, Daitoku
;     TITLE OF INVENTION:  Mariba and Human Bax-Dependent
;     TITLE OF INVENTION:  MYC-Interacting Protein (BIN1) and Uses Therefor
;     NUMBER OF SEQUENCES:  7
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Howson and Howson
;       STREET:  Spring House Corporate Cntr, P.O Box 457
;       CITY:  Spring House
;       STATE:  Pennsylvania
;       COUNTRY:  USA
;       ZIP:  19477
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.30
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/652,972A
;       FILING DATE:  24-MAY-1996
;       CLASSIFICATION:  514
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/435,454
;       FILING DATE:  05-MAY-1995
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Bak, Mary E.
;       REGISTRATION NUMBER:  31,215
;       REFERENCE/DOCKET NUMBER:  WS1608USA
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  215-540-9200
;       TELEFAX:  215-540-5818
;   INFORMATION FOR SEQ ID NO:  4:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  451 amino acids
;       TYPE:  amino acid
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  protein
US-08-652-972A-4


  Query Match             65.1%;  Score 28;  DB 2;  Length 451;
  Best Local Similarity   85.7%;  Pred. No. 2e+02;
  Matches    6;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       5 KIASNIQ 9
            ||||| |
Db      10 KIASNVQ 16


RESULT  12
US-08-919-145-6
; Sequence 6, Application US/08919145
; Patent No. 5958753
;  GENERAL INFORMATION:
;     APPLICANT:  Prendergast, George C.
;     TITLE OF INVENTION:  Bax, A Bcl2 Interacting Protein, and
;     TITLE OF INVENTION:  Uses Therefor
;     NUMBER OF SEQUENCES:  8
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Howson and Howson
;       STREET:  Spring House Corporate Cntr, P.O.  Box 457
;       CITY:  Spring House
;       STATE:  Pennsylvania
;       COUNTRY:  USA
;       ZIP:  19477
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.30
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/919,145
;       FILING DATE:
;       CLASSIFICATION:  514
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 60/025,482
;       FILING DATE:  29-AUG-1996
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Kodroff, Cathy A.
;       REGISTRATION NUMBER:  33,980
;       REFERENCE/DOCKET NUMBER:  WS173AUSA
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  215-540-9200
;       TELEFAX:  215-540-5818
;   INFORMATION FOR SEQ ID NO:  6:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  451 amino acids
;       TYPE:  amino acid
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  protein
US-08-919-145-6


  Query Match             65.1%;  Score 28;  DB 2;  Length 451;
  Best Local Similarity   85.7%;  Pred. No. 2e+02;
  Matches    6;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       5 KIASNIQ 9
            ||||| |
Db      10 KIASNVQ 16


RESULT  13
US-08-870,126-4
; Sequence 4, Application US/08870126
; Patent No. 6048702
;  GENERAL INFORMATION:
;     APPLICANT:  Prendergast, George C.
```

```
;       APPLICANT:  Sakamuro, Dorrtoko
;       TITLE OF INVENTION:  Murine and Human Box Dependent
;       TITLE OF INVENTION:  Myc Interacting Protein (BIN1) and Uses Therefor
;       NUMBER OF SEQUENCES:  14
;       CORRESPONDENCE ADDRESS:
;         ADDRESSEE:  Howson and Howson
;         STREET:  Spring House Corporate Cntr, P.O. Box 457
;         CITY:  Spring House
;         STATE:  Pennsylvania
;         COUNTRY:  USA
;         ZIP:  19477
;       COMPUTER READABLE FORM:
;         MEDIUM TYPE:  Floppy disk
;         COMPUTER:  IBM PC compatible
;         OPERATING SYSTEM:  PC-DOS/MS-DOS
;         SOFTWARE:  Patentin Release #1.0, Version #1.30
;       CURRENT APPLICATION DATA:
;         APPLICATION NUMBER:  US/08/439,454
;         FILING DATE:  05-MAY-1995
;       PRIOR APPLICATION DATA:
;         APPLICATION NUMBER:  US 08/452,972
;         FILING DATE:  24-MAY-1995
;       ATTORNEY/AGENT INFORMATION:
;         NAME:  Kodroff, Cathy A.
;         REGISTRATION NUMBER:  33,980
;         REFERENCE/DOCKET NUMBER:  WST60CUSA
;       TELECOMMUNICATION INFORMATION:
;         TELEPHONE:  215-540-9200
;         TELEFAX:  215-540-5818
;     INFORMATION FOR SEQ ID NO:  4:
;       SEQUENCE CHARACTERISTICS:
;         LENGTH:  451 amino acids
;         TYPE:  amino acid
;         TOPOLOGY:  linear
;       MOLECULE TYPE:  protein
US-09-744-282-4

  Query Match          65.1%;  Score 28;  DB 4;  Length 451;
  Best Local Similarity   85.7%;  Pred. No. 2e-02;
  Matches    6;  Conservative    0;  Mismatches    1;  Indels    1;  Gaps    0;

Qy        3 KIASNIQ 9
            |||||| |
Db      170 KIASNVQ 176


RESULT  14
US-08-744-889-6
; Sequence 6, Application US/08744889
; Patent No. 5714459
; GENERAL INFORMATION:
;       APPLICANT:  Prendergast, George C.
;       TITLE OF INVENTION:  Bax, A Bin1 Interacting Protein, and
;       TITLE OF INVENTION:  Uses Therefor
;       NUMBER OF SEQUENCES:  8
;       CORRESPONDENCE ADDRESS:
;         ADDRESSEE:  Howson and Howson
;         STREET:  Spring House Corporate Cntr, P.O. Box 457
;         CITY:  Spring House
;         STATE:  Pennsylvania
;         COUNTRY:  USA
;         ZIP:  19477
;       COMPUTER READABLE FORM:
;         MEDIUM TYPE:  Floppy disk
;         COMPUTER:  IBM PC compatible
;         OPERATING SYSTEM:  PC-DOS/MS-DOS
;         SOFTWARE:  Patentin Release #1.0, Version #1.30
;       CURRENT APPLICATION DATA:
;         APPLICATION NUMBER:  US/08/744,889
;         FILING DATE:
;         CLASSIFICATION:
;       PRIOR APPLICATION DATA:
;         APPLICATION NUMBER:  US/08/919,145
;         FILING DATE:
;         APPLICATION NUMBER:  US 60/025,482
;         FILING DATE:  29-AUG-1996
;       ATTORNEY/AGENT INFORMATION:
;         NAME:  Kodroff, Cathy A.
;         REGISTRATION NUMBER:  33,980
;         REFERENCE/DOCKET NUMBER:  WST73AUSA
;       TELECOMMUNICATION INFORMATION:
;         TELEPHONE:  215-540-9200
;         TELEFAX:  215-540-5818
;     INFORMATION FOR SEQ ID NO:  6:
;       SEQUENCE CHARACTERISTICS:
;         LENGTH:  451 amino acids
;         TYPE:  amino acid
;         TOPOLOGY:  linear
;       MOLECULE TYPE:  protein
US-09-744-889-6


  Query Match          65.1%;  Score 28;  DB 4;  Length 451;
  Best Local Similarity   85.7%;  Pred. No. 2e-02;
  Matches    6;  Conservative    0;  Mismatches    1;  Indels    1;  Gaps    0;

Qy        3 KIASNIQ 9
            |||||| |
Db      170 KIASNVQ 176


RESULT  15
PCT-US96-06231A-4
; Sequence 4, Application PC/US9606231A
; GENERAL INFORMATION:
;       APPLICANT:  Wistar Institute of Anatomy &, Biology
;       TITLE OF INVENTION:  Murine and Human Box Dependent
;       TITLE OF INVENTION:  Myc Interacting Protein (BIN1) and Uses Therefor
;       NUMBER OF SEQUENCES:  7
;       CORRESPONDENCE ADDRESS:
;         ADDRESSEE:  Howson and Howson
;         STREET:  Spring House Corporate Cntr, P.O. Box 457
;         CITY:  Spring House
;         STATE:  Pennsylvania
;         COUNTRY:  USA
;         ZIP:  19477
;       COMPUTER READABLE FORM:
;         MEDIUM TYPE:  Floppy disk
;         COMPUTER:  IBM PC compatible
;         OPERATING SYSTEM:  PC-DOS/MS-DOS
;         SOFTWARE:  Patentin Release #1.0, Version #1.30
;       CURRENT APPLICATION DATA:
;         APPLICATION NUMBER:  PCT/US96/06231A
;         FILING DATE:
;         CLASSIFICATION:
;       PRIOR APPLICATION DATA:
;         APPLICATION NUMBER:  US 08/435,454
;         FILING DATE:  05-MAY-1995
;       ATTORNEY/AGENT INFORMATION:
;         NAME:  Bak, Mary E.
;         REGISTRATION NUMBER:  31,215
;         REFERENCE/DOCKET NUMBER:  WS160APCT
;       TELECOMMUNICATION INFORMATION:
;         TELEPHONE:  215-540-9200
;         TELEFAX:  215-540-5818
;     INFORMATION FOR SEQ ID NO:  4:
;       SEQUENCE CHARACTERISTICS:
;         LENGTH:  451 amino acids
;         TYPE:  amino acid
;         TOPOLOGY:  linear
;       MOLECULE TYPE:  protein
```

PCT-US96-09231A-4

Query Match 65.1%; Score 28; DB 5; Length 451;
Best Local Similarity 85.7%; Prod. No. 2e-02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 3 KIASNI 9
||:|:|
Db 1 KIASNV 16

Search completed: August 26, 2002, 06:24:28
Job time: 86.1 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 26, 2002, 06:10:48 ; Search time 80.5 Seconds
(without alignments)
15,518 Million cell updates/sec

Title: US-09-744-282-5
Perfect score: 65
Sequence: 1 PIFGRSFTLASSE 13

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 284138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 284138

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_71:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	65	100.0	383	2	A49562	cartilage glycopro
2	62	95.4	405	2	S61551	breast-regressing
3	61	93.8	383	2	S51327	heparin-binding gl
4	45	69.2	525	2	T44145	chitinase (EC 3.2.
5	41	63.1	203	2	C25511	Ne protein - trout
6	40	61.5	415	2	C87328	conserved hypothet
7	40	61.5	554	2	A56596	chitinase (EC 3.2.
8	40	61.5	1635	2	T14075	chitinase (EC 3.2.
9	39	60.0	113	2	I51025	transcription fact
10	39	60.0	483	2	A53918	chitinase (EC 3.2.
11	39	60.0	1199	2	A40670	nuclear envelope p
12	38	58.5	567	2	I49942	hypothetical prote
13	37	56.9	177	1	A69473	conserved hypothet
14	37	56.9	205	2	G69379	conserved hypothet
15	37	56.9	610	2	I49485	prostate cystatin
16	37	56.9	641	2	I05497	hypothetical prote
17	37	56.9	684	2	B69308	conserved hypothet
18	37	56.9	1074	2	F72217	conserved hypothet
19	36	55.4	462	2	B70593	hypothetical prote
20	36	55.4	462	2	A75425	pir:p:pyrimidin
21	36	55.4	617	2	I15408	hypothetical prote
22	36	55.4	654	2	I38605	oviductal glycopro
23	36	55.4	676	2	AE2417	hypothetical prote
24	36	55.4	688	2	E88197	protein him-14 [im
25	36	55.4	765	2	E84553	hypothetical prote
26	35	53.8	175	2	H97174	uncharacterized pr
27	35	53.8	179	2	AG1994	hypothetical prote
28	35	53.8	364	2	C71295	probable octapreny
29	35	53.8	386	2	AG0081	conserved hypothet
30	35	53.8	387	2	S56411	hypothetical 47K p
31	35	53.8	387	2	B91274	probable synthetas
32	35	53.8	387	2	B86115	probable synthetas
33	35	53.8	387	2	AD0890	conserved hypothet
34	35	53.8	387	2	AB1650	conserved hypothet
35	35	53.8	402	2	AD3373	NAD(FAD) utilizing
36	35	53.8	444	2	B85725	hypothetical prote
37	35	53.8	527	2	I15606	hypothetical prote
38	35	53.8	537	2	S57197	oviduct-specific g
39	35	53.8	539	2	I40470	estrogen-dependent
40	35	53.8	550	2	I01776	hypothetical prote
41	35	53.8	591	2	S74790	hypothetical prote
42	35	53.8	849	2	B89430	70 kDa protein - A
43	35	53.8	1148	2	S51855	hypothetical prote
44	35	53.8	3716	2	E29099	probable ATP-dep pr
45	34	52.3	208	2	DB3895	hypothetical prote

ALIGNMENTS

RESULT 1
A49562
cartilage glycoprotein gp39 precursor - human
N;Alternate names: 39K synovial protein
C;Species: Homo sapiens (man)
C;Date: 25-Mar-1995 #sequence_revision 25-Mar-1995 #text_change 21-Jul-1995
C;Accession: A49562; S10677; A33162
R;Hakala, B.E.; White, C.; Recklies, A.D.
J. Biol. Chem. 268, 25803-25810, 1993
A;Title: Human cartilage gp-39, a major secretory product of articular chondrocytes a
A;Reference number: A49562; MUID:94064658
A;Accession: A49562
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-383 <HAK>
A;Cross-references: GB:M80927; NID:g348611; PIDN:AAA16074.1; PID:g348611
R;Nyirkos, P.; Golds, E.E.
Biochem. J. 269, 265-268, 1990
A;Title: Human synovial cells secrete a 39 kDa protein similar to a bovine mammary pr
A;Reference number: S10677; MUID:90328083
A;Accession: S10677
A;Molecule type: protein
A;Residues: 22-40,'X',42-45 <NY2>
C;Superfamily: Streptomyces chitinase chi40
C;Keywords: cartilage; extracellular protein; glycoprotein
F;1-21/Domain: signal sequence #status predicted <SIG>
F;22-383/Product: cartilage glycoprotein gp39 #status predicted <MAT>

Query Match 100.0%; Score 65; DB 2; Length 383;
Best Local Similarity 100.0%; Pred. No. 0.00011;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PIFGRSFTLASSE 13
|||||||||||||
Db 259 PIFGRSFTLASSE 271

RESULT 2
S61551
breast-regressing protein brp39 precursor - mouse
C;Species: Mus musculus (house mouse)
C;Date: 19-Mar-1997 #sequence_revision 25-Apr-1997 #text_change 20-Jun-2000
C;Accession: S61551; S61550; I48271
R;Morrison, B.W.; Leder, P.
Oncogene 9, 3417-3426, 1994
A;Title: neu and ras initiate murine mammary tumors that share genetic markers detera
A;Reference number: I48271; MUID:95060797
A;Accession: S61551
A;Status: nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1-405 <MOR1>

A;Title: Sequence and expression of the Cc gene, a member of the dopa decarboxylase g
A;Reference number: A93634; MUID:86312887
A;Accession: C25511
A;Molecule type: DNA
A;Residues: 1-203 <EVE>
A;Cross-references: GB:X04227; NID:g7695; PIDN:CAA27807.1; PID:g295743
C;Genetics:
A;Gene: FlyBase:l(2)37Cc
A;Cross-references: FlyBase:FBgn0002031
A;Introns: 67/3

  Query Match 63.1%; Score 41; DB 2; Length 203;
  Best Local Similarity 88.9%; Pred. No. 2.6;
  Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 IFGRSFILA 10
 |||| ||||
Db 97 IFGREFILA 105

RESULT 6
C87428
conserved hypothetical protein CC0638 [imported] - Caulobacter crescentus
C;Species: Caulobacter crescentus
C;Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 20-Apr-2001
C;Accession: C87428
R;Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg,
 B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Hatt, J.M.; Ko
 to, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A;Title: Complete Genome Sequence of Caulobacter crescentus.
A;Reference number: A87249; MUID:21173698; PMID:11259647
A;Accession: C87428
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-415 <SIO>
A;Cross-references: GB:AE006057; NID:g13421851; PIDN:AAK22234.1; GSPDB:GN01148
C;Genetics:
A;Gene: CC0638

  Query Match 61.5%; Score 40; DB 2; Length 415;
  Best Local Similarity 70.0%; Pred. No. 8.7;
  Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 PIFGRSFILA 10
 |||| | :|
Db 91 PIFGRKFLMA 60

RESULT 7
A56596
chitinase (EC 3.2.1.14) - tobacco hornworm
C;Species: Manduca sexta (tobacco hornworm)
C;Date: 11-Aug-1995 #sequence_revision 11-Aug-1995 #text_change 26-Jan-1996
C;Accession: A56596

R;Kramer, K.J.; Corpuz, L.; Choi, H.K.; Muthukrishnan, S.
Insect Biochem. Mol. Biol. 23, 691-701, 1993
A;Title: Sequence of a cDNA and expression of the gene encoding epidermal and gut chitin
A;Reference number: A56596; MUID:93357793
A;Accession: A56596
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-554 <KRA>
A;Cross-references: GB:U02270; GB:S64757; NID:g406048; PID:g406049
A;Experimental source: larvae
A;Note: Species extracted from NCBI backbone (NCBIN:19417, NCBIP:19418)
C;Keywords: glycosidase; hydrolase; polysaccharide degradation

Query Match 61.5%; Score 40; DB 2; Length 554;
Best Local Similarity 68.6%; Pred. No. 1.;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PIFGRSFILAS 11
| :|:|:|:::
Db 268 PFYGKSFILSA 278

RESULT 8
I14075
chitinase (EC 3.2.1.14) - yellow fever mosquito
C;Species: Aedes aegypti (yellow fever mosquito)
C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 2-Sep-1999
C;Accession: I14075
R;de la Vega, H.; Specht, C.A.; Liu, Y.; Robbins, P.W.
Insect Mol. Biol. 7, 233-239, 1997
A;Title: Chitinases are a multi-gene family in Aedes, Anopheles, and Drosophila.
A;Reference number: Z17872
A;Accession: I14075
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-1645 <DEL>
A;Cross-references: EMBL:AF026492; NID:g2564720; PID:g2564721; PIDN:AAB61850.1
C;Genetics:
A;Gene: CHI2
A;Cross-ref: [illegible]
C;Keywords: glycosidase; hydrolase; polysaccharide degradation

Query Match 61.5%; Score 40; DB 2; Length 1645;
Best Local Similarity 54.8%; Pred. No. 48;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 1 PIFGRSFILASSE 13
| :|:|||:|| :
Db 484 PLYGQSFILADQK 496

RESULT 9
I51025
transcription factor - eastern newt (fragment)
C;Species: Notophthalmus viridescens, Triturus viridescens (eastern newt)
C;Date: 18-Sep-1996 #sequence_revision 18-Sep-1996 #text_change 24-Sep-1999
C;Accession: I51025
R;Simon, H.-G.; Nelson, C.; Gotz, D.; Lanier, E.; Martin, B.A.; Labib, C.
Dev. Dyn. 202, 1-12, 1995
A;Title: Differential expression of myogenic regulatory genes and Msx 1 during dedifferen
A;Reference number: I51024; MUID:95218226
A;Accession: I51025
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-114 <SIM>
A;Cross-references: EMBL:X82837; NID:g577786; PIDN:CAA58044.1; PID:g577787
C;Genetics:
A;Gene: Msx 1
C;Superfamily: unassigned homeobox proteins; homeobox homology
C;Keywords: DNA binding; homeobox; nucleus; transcription regulation
F;1-45/Domain: homeobox homology (fragment) <HOX>

Query Match 60.0%; Score 39; DB 2; Length 1,3;
Best Local Similarity 72.7%; Pred. No. 3.;
Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 PIFGRSFILAS 11
|||| || | |
Db 64 PIFGLSFPIGS 74

RESULT 10
A53918
chitinase (EC 3.2.1.14) precursor - braconid wasp (Cheloninae sp.)
C;Species: Cheloninae sp.
C;Date: 28-Jul-1995 #sequence_revision 28-Jul-1995 #text_change 2-Jul-2000
C;Accession: A53918
R;Krishnan, A.; Nair, P.N.; Jones, D.
J. Biol. Chem. 269, 20971-20976, 1994
A;Title: Isolation, cloning, and characterization of new chitinase stored in active f
A;Reference number: A53918; MUID:94342256
A;Accession: A53918
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-483 <KRI>
A;Cross-references: GB:U10422; NID:g544534; PIDN:AAA61444.1; PID:g544535
C;Keywords: glycosidase; hydrolase; polysaccharide degradation

Query Match 60.0%; Score 39; DB 2; Length 483;
Best Local Similarity 54.5%; Pred. No. 19;
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PIFGRSFILAS 11
| :|:|||::
Db 268 PAYGKSFILSN 278

RESULT 11
A40670
n, ocular envelope protein P0M 121 - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 03-May-1994 #sequence_revision 03-May-1994 #text_change 06-Nov-1999
C;Accession: A40670
R;Hallberg, E.; Wozniak, R.W.; Blobel, G.
J. Cell Biol. 122, 513-521, 1993
A;Title: An integral membrane protein of the pore membrane domain of the nuclear enve
A;Reference number: A40670; MUID:93328754
A;Accession: A40670
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-1199 <HAL>
A;Cross-references: GB:Z21913; NID:g396746; PIDN:CAA79725.1; PID:g396747
F;804-807,845-849,956-960,1010-1014,1047-1051/Other: pentapeptide motif (X

Query Match 60.0%; Score 39; DB 2; Length 1199;
Best Local Similarity 58.3%; Pred. No. 4.;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 PIFGRSFILASS 12
|:|| ||| :|
Db 951 PSFGSSFIPGNS 962

RESULT 12
I49942
hypothetical protein F17I14.250 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 02-Jun-2
C;Accession: I49942
R;Bevan, M.; Hilbert, H.; Bram, M.; Hilson, P.; Brandt, A.; Unseld, M.; Lemcke
submitted to the Protein Sequence Database; April 2000

A;Reference number: Z24490
A;Accession no.: I49942
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-567 <REV>
A;Cross-references: EMBL:AL033994; GSPDB:GN00063; AISP:EF7114_2
A;Experimental source: cultivar Columbia; BAC clone EF7114
C;Genetics:
A;Gene: AISP:EF7114.250
A;Map position: 5
A;Introns: 54/3; 294/3; 403/3; 452/3

Query Match 56.5%; Score 37; DB 2; Length 567;
Best Local Similarity 77.8%; Pred. No. 90;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PIFGRSFIL 9
|:||:|:|
Db 142 PSFGRDFAIA 151

RESULT 15
T39485
probable cystathionine gamma-synthase - fission yeast (Schizosaccharomyces pombe)
C;Species: Schizosaccharomyces pombe
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
C;Accession: T39485
R;Lyne, M.; Rajandream, M.A.; Barrell, B.G.; Lucas, M.; Gaillardin, C.
submitted to the EMBL Data Library, August 1997
A;Reference number: Z21858
A;Accession: T39485
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-610 <LYN>
A;Cross-references: EMBL:AL031439; PIDN:CAA20484.1; GSPDB:GN00063; SPDB:SPBC1604.09c
A;Experimental source: strain 972h-; cosmid c1604
C;Genetics:
A;Gene: SPDB:SPBC1604.09c
A;Map position: 2

Query Match 56.9%; Score 37; DB 2; Length 610;
Best Local Similarity 63.6%; Pred. No. 51;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 PIFGRSFILAS 11
|:|:|||||
Db 555 PSLGINFTLAS 565

Search completed: August 26, 2002, 06:10:50
Job time: 111 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 26, 2002, 06:29:04 ; Search time 35.06 Seconds
(without alignments)
13,986 Million cell updates/sec

Title: US-09-744-282-5
Perfect score: 65
Sequence: 1 PTPGRSFILASSE 13

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt 40:*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% query Match	Length	DB	ID	Description
1	65	100.0	383	1	C3L1_HUMAN	P36222, homo sapien
2	62	95.4	381	1	C3L1_MOUSE	Q61553 mus musculu
3	54	83.1	390	1	C3L2_HUMAN	Q15782, homo sapien
4	41	63.1	203	1	L2CC_DROME	P14199 drosophila
5	41	63.1	554	1	CHIT_MANSE	P36382 manduca sex
6	39	60.0	1199	1	NI21_RAT	P52594 rattus norv
7	39	60.0	1229	1	NI21_HUMAN	Q92820 homo sapien
8	37	56.9	177	1	YH86_ARCFU	O28888 archaeoglob
9	37	56.9	818	1	CDB1_HUMAN	Q92613 homo sapien
10	36	55.4	527	1	OGP_PIG	Q28950 sus scrofa
11	36	55.4	617	1	CHIT_CAEEL	Q11174 caenorhabdi
12	36	55.4	623	1	OGP_PAPAN	P49711 papio anubi
13	36	55.4	678	1	OGP_HUMAN	Q13489 homo sapien
14	36	55.4	1997	1	OTOF_HUMAN	Q96IH1 homo sapien
15	36	55.4	1997	1	OTOF_MOUSE	Q9QSI1 mus musculu
16	35	53.8	387	1	YJFC_ECOLI	P33222 escherichia
17	35	53.8	522	1	OGP_BOVIN	Q28162 bos taurus
18	35	53.8	539	1	OGP_SHEEP	Q28434 ovis aries
19	35	53.8	592	1	YD72_MYCPN	P75472 mycoplasma
20	35	53.8	721	1	OGP_MOUSE	Q99LI0 mus musculu
21	35	53.8	747	1	ORPP_HUMAN	Q96RI4 homo sapien
22	34	52.3	267	1	YTHC_XANAX	Q2P1F4 xanthomonas
23	34	52.3	435	1	HISX_BUCAI	P57211 buchnera ap
24	34	52.3	454	1	ARX_BRARE	Q42315 brachydanio
25	34	52.3	477	1	INGR_MOUSE	P17261 mus musculu
26	34	52.3	695	1	CMC1_DROME	Q9VA78 drosophila
27	34	52.3	798	1	CDBE_HUMAN	Q96TC9 homo sapien
28	34	52.3	1012	1	POLG_IBDV6	Q27Q2 avian infec
29	34	52.3	1159	1	DF3A_VIBCH	P52722 vibrio chol
30	34	52.3	1190	1	ZO2_HUMAN	Q92802 homo sapien
31	34	52.3	1306	1	MSB2_YEAST	P53334 saccharomyc
32	34	52.3	1385	1	FAI1_SCHPO	P78756 schizosacch
33	33	50.8	120	1	RET3_CHICK	P49220 gallus gall
34	33	50.8	129	1	AA3R_MOUSE	Q61618 mus musculu
35	33	50.8	255	1	CIB3_HUMAN	Q09472 homo sapien
36	33	50.8	258	1	EST2_USEPL	Q73547 pseudomonas
37	33	50.8	224	1	SSAK_SALTY	P74853 salmonella
38	33	50.8	268	1	TRPA_HAEIN	P44759 haemophilus
39	33	50.8	304	1	NOL1_KRIS4	Q12865 kluyveromyc
40	33	50.8	351	1	VGRR_METTH	Q50891 methanobact
41	33	50.8	466	1	SYN_VIBCH	Q9KS14 vibrio cho
42	33	50.8	488	1	SUCE_AGRVI	P33918 agrobacteri
43	33	50.8	542	1	MIT7_NEUCR	P38675 neurospora
44	33	50.8	555	1	NFL_GOLJA	Q25167 saturnix sa
45	33	50.8	566	1	MXIB_SHIFL	Q41641 shigella fl

ALIGNMENTS

RESULT 1
C3L1_HUMAN
ID C3L1_HUMAN STANDARD; PRT: 383 AA.
AC P36222; P30923;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Chitinase-3 like protein 1 precursor (Cartilage glycoprotein-39) (GP
DE 39) (39 kDa synovial protein) (YKL-40).
GN CHI3L1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC TISSUE=Cartilage;
RX MEDLINE=94064658; PubMed=8245017;
RA Hakala B.E., White C., Recklies A.D.;
RT "Human cartilage gp-39, a major secretory product of articular
RT chondrocytes and synovial cells, is a mammalian member of a chitinase
RT protein family.";
RL J. Biol. Chem. 268:25803-25810(1993).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Blood;
RX MEDLINE=97386591; PubMed=9244440;
RA Kelly M., Krause S.W., Andresson K.;
RT "Molecular characterization of the gene for human cartilage gp-39
RT (CHI3L1), a member of the chitinase protein family and marker for
RT late stages of macrophage differentiation.";
RL Genomics 43:221-225(1997).
RN [3]
RP SEQUENCE OF 22-45.
RX MEDLINE=90328983; PubMed=2375255;
RA Nyirkos P., Golds E.E.;
RT "Human synovial cells secrete a 39 kDa protein similar to a bovine
RT mammary protein expressed during the non-lactating period.";
RL Biochem. J. 267:265-268(1990).
CC -!- FUNCTION: MAY PLAY AN IMPORTANT ROLE IN THE CAPACITY OF CELLS TO
CC RESPOND TO AND COPE WITH CHANGES IN THEIR ENVIRONMENT.
CC -!- SUBCELLULAR LOCATION: Extracellular.
CC -!- TISSUE SPECIFICITY: PRESENT IN ARTICULAR CHONDROCYTES, SYN VIAL
CC CELLS AS WELL AS IN LIVER, NEUTROPHILS IN SYN. TISSUES, LUNG,
CC PANCREAS, MONONUCLEAR CELLS, OSTEOBLASTS.
CC -!- PTM: GLYCOSYLATED.
CC -!- SIMILARITY: BELONGS TO FAMILY 18 (CHITINASE).
CC ---------- -- -
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is not modified
CC and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch)
CC ---------- -- -

```
ID   EMBL: M[illegible] standard; mRNA; AAA19074.1; -.
AC   EMBL: Y[illegible] AAA96661.1; -.
DR   EMBL; Y[illegible] AAA96661.1; JOINED.
DR   EMBL; Y[illegible] AAA96661.1; JOINED.
DR   EMBL; Y[illegible] AAA96661.1; JOINED.
DR   EMBL; Y[illegible] AAA96661.1; JOINED.
DR   [illegible] 21[illegible]
DR   P[illegible] A33162; A33162.
DR   MIM; [illegible]
DR   InterPro; IPR001579; Chitinase_2.
DR   InterPro; IPR001223; Glyco_hydro_18.
DR   Pfam; PF[illegible]; Glyco_hydro_18; 1.
DR   PROSITE; PS51095; CHITINASE_18; FALSE_NEG.
KW   Hydrolase; Signal.
FT   SIGNAL       [illegible]      21
FT   CHAIN        22     383       CHITINASE 3 LIKE PROTEIN 1.
FT   CARBOHYD     [illegible]
SQ   SEQUENCE   [illegible] MW; [illegible] CRC64;


  Query Match             [illegible]; Score [illegible]; DB 1; Length [illegible];
  Best Local Similarity   [illegible]; Pred. No. [illegible];
  Matches  [illegible]; Conservative  [illegible]; Mismatches  [illegible]; Indels  [illegible]; Gaps  [illegible];

Qy       1 FIPGKSFILASSE 13
             [illegible]
Db     260 FIPGKSFILASSE 272


RESULT [illegible]
[illegible]_HUMAN
ID   [illegible] STANDARD;      PRT:   [illegible] AA.
AC   [illegible];
DT   [illegible] (Rel. 40, Created)
DT   [illegible] (Rel. 40, Last sequence update)
DT   [illegible] (Rel. 40, Last annotation update)
DE   [illegible] precursor ([illegible]) (Chondrocyte protein
DE   [illegible]).
GN   [illegible]
OS   [illegible] (Human).
OC   [illegible]; Metazoa; Chordata; [illegible]; Vertebrata; [illegible];
OC   [illegible]; Primates; [illegible]; [illegible]; [illegible];
OX   [illegible];
RN   [1]
RP   SEQUENCE FROM N.A. ([illegible]).
RA   [illegible]
RA   [illegible];
RT   "[illegible]
RT   [illegible]
RL   [illegible]
RN   [2]
RP   SEQUENCE FROM N.A. (SHORT ISOFORM).
RC   TISSUE=Articular cartilage;
RX   [illegible];
RA   [illegible]
RT   "[illegible]
RT   [illegible]
RL   [illegible]
CC   -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS; A SHORT FORM (SHOWN HERE) AND A
CC       LONG FORM; MAY BE PRODUCED BY ALTERNATIVE SPLICING.
CC   -!- TISSUE SPECIFICITY: HIGHEST EXPRESSION IN [illegible]
CC       BY SYNOVIOCYTES, LUNG AND HEART. NOT DETECTED IN BRAIN, SPLEEN,
CC       PANCREAS, AND LIVER.
CC   -!- SIMILARITY: BELONGS TO FAMILY 18 OF GLYCOSYL HYDROLASES.
CC   -----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://[illegible]
CC   or send an email to license@isb-sib.ch).
CC   -----------------------------------------------------------------------
DR   EMBL; [illegible]; AAA[illegible].1; -.
DR   EMBL; [illegible]; AAA[illegible].1; -.
DR   EMBL; [illegible]; AAA[illegible].1; ALT_INIT.
DR   MIM; [illegible]; -.
DR   InterPro; IPR[illegible]; Chitinase_2.
DR   InterPro; IPR001223; Glyco_hydro_18.
DR   Pfam; PF00704; Glyco_hydro_18; 1.
DR   PROSITE; PS51095; CHITINASE_18; FALSE_NEG.
KW   Hydrolase; Signal; Alternative splicing.
FT   SIGNAL       1     26       POTENTIAL.
FT   CHAIN       27    390       CHITINASE 3 LIKE PROTEIN 2.
FT   CARBOHYD    35     35       N-LINKED (GlcNAc...) (POTENTIAL).
FT   VARSPLIC     1     12       MGAIIMDGKSLW -> SQQLWLGSPDEVWLYLVCPEHIHFT
FT                               AIWGSKIKFGSIDFSGHEIS (IN LONG ISOFORM).
SQ   SEQUENCE   390 AA;  43500 MW;  97B86A2F3AA[illegible] CRC64;
```

Query Match 83.1%; Score 54; DB 1; Length 3??;
Best Local Similarity 76.9%; Pred. No. 0.0059;
Matches 10; Conservative 2; Mismatches 1; Indels 1; Gaps 0;

Qy 1 PTFGRSFTLASSE 13
|| | | | | | | | | | | | | |
Db 267 PTYGRSFTLASAE 279

RESULT 4
L20C_DROME
ID L20C_DROME STANDARD; PRT; 283 AA.
AC P24156;
DT 01-MAR-1993 (Rel. 25, Created)
DT 01-MAR-1993 (Rel. 25, Last sequence update)
DT 01-???-1999 (Rel. 36, Last annotation update)
DE L(2)37Cc protein.
GN L(2)37Cc.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE 86312887; PubMed 3092183;
RA Eveleth D.D. Jr., Marsh J.L.;
RT "Sequence and expression of the Cc gene, a member of the dopa
RT decarboxylase gene cluster of Drosophila: possible translational
RT regulation.";
RL Nucleic Acids Res. 14:6161-6183(1986).
CC -!- FUNCTION: REQUIRED FOR LARVAL METABOLISM OR FOR THE PROGRESSION
CC OF THE LARVA INTO A PUPA.
CC -!- SIMILARITY: BELONGS TO THE PROHIBITIN FAMILY.
CC ---------- ----------
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC ----------
DR EMBL; X04228; CAA27810.1; -.
DR EMBL; X04227; CAA27807.1; -.
DR PIR; S25511; S25511.
DR FlyBase; FBgn0002043; l(2)37Cc.
DR InterPro; IPR001107; Band_7.
DR Pfam; PF01145; Band_7; 1.
DR SMART; SM00244; PHB; 1.
KW Developmental protein.
SQ SEQUENCE 283 AA; 32817 MW; B15D0B5D03B62A11 CRC64;

Query Match 63.1%; Score 41; DB 1; Length 283;
Best Local Similarity 88.9%; Pred. No. 0.95;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 TFGRSFTLA 10
|||| | ||
Db 67 TFGRFFTLA 105

RESULT 5
CHIT_MANSE
ID CHIT_MANSE STANDARD; PRT; 554 AA.
AC P36362;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 01-FEB-1995 (Rel. 31, Last annotation update)
DE Endochitinase precursor (EC 3.2.1.14).
OS Manduca sexta (Tobacco hawkmoth) (Tobacco hornworm).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
OC Sphingiodea; Sphingidae; Sphinginae; Manduca.
OX NCBI_TaxID=7130;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE 93357793; PubMed 8354525;
RA Kramer K.J., Corpuz L., Choi H.K., Muthukrishnan S.;
RT "Sequence of a cDNA and expression of the gene encoding epidermal and
RT gut chitinases of Manduca sexta.";
RL Insect Biochem. Mol. Biol. 23:691-701(1993).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE 97215580; PubMed 9061927;
RA Choi H.K., Choi K.H., Kramer K.J., Muthukrishnan S.;
RT "Isolation and characterization of a genomic clone for the gene of an
RT insect molting enzyme, chitinase.";
RL Insect Biochem. Mol. Biol. 27:37-47(1997).
CC -!- FUNCTION: DIGEST CHITIN IN THE EXOSKELETON DURING THE MOLTING
CC PROCESS.
CC -!- CATALYTIC ACTIVITY: Hydrolysis of the 1,4-beta-linkages of N-
CC acetyl-D-glucosamine polymers of chitin.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: EPIDERMIS AND GUT.
CC -!- DEVELOPMENTAL STAGE: HIGH LEVELS SEEN IN THE EPIDERMIS ON DAY 0,
CC BUT RAPIDLY DISAPPEARS AND IS UNDETECTED ON DAYS 1-4 OF FIFTH
CC INSTAR. IT REAPPEARS ON DAY 5 AND PEAKS ON DAY 7 AFTER WHICH A
CC RAPID DECLINE IS SEEN. IN THE GUT IT DISAPPEARS AFTER DAY 0 WITH LOWER
CC LEVELS SEEN ON DAYS 0, 7 AND 8.
CC -!- SIMILARITY: BELONGS TO CHITINASE CLASS II (FAMILY 18 OF GLYCOSYL-
CC HYDROLASES).
CC ---------- ----------
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC ----------
DR EMBL; U02270; AAC04924.1; -.
DR EMBL; L17274; AAB59682.1; -.
DR InterPro; IPR002557; Chitin_Binding.
DR InterPro; IPR001579; Chitinase_2.
DR InterPro; IPR001223; Glyco_hydro_18.
DR Pfam; PF01607; Chitin_bind_2; 1.
DR Pfam; PF00704; Glyco_hydro_18; 1.
DR SMART; SM00494; ChtBD2; 1.
DR PROSITE; PS01095; CHITINASE_18; 1.
KW Hydrolase; Glycosidase; Chitin degradation; Signal; Glycoprotein.
FT SIGNAL 1 19 POTENTIAL.
FT CHAIN 20 554 ENDOCHITINASE.
FT DOMAIN 396 453 SER/THR-RICH.
FT ACT_SITE 146 146 PROTON DONOR (BY SIMILARITY).
FT CARBOHYD 85 85 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 303 303 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 467 467 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 545 545 N-LINKED (GLCNAC...) (POTENTIAL).
SQ SEQUENCE 554 AA; 62203 MW; 39B9D756C96C3490 CRC64;

Query Match 61.5%; Score 40; DB 1; Length ?;
Best Local Similarity 63.6%; Pred. No. 4.2;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PTFGRSFTLAS 11
| : |||||::
Db 268 PFYGRSFTLSA 278

RESULT 6
NI2L_RAT

```
ID   N[illegible]_RAT       STANDARD;      PRT;  1199 AA.
AC   [illegible]
DT   [illegible] (Rel. 34, Created)
DT   [illegible] (Rel. 34, Last sequence update)
DT   [illegible] (Rel. 34, Last annotation update)
DE   Nuclear envelope pore membrane protein POM 121 (Pore membrane protein
DE   of [illegible] kDa) (P145).
GN   N[illegible]
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID [illegible];
RN   [1]
RP   SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC   STRAIN=SPRAGUE-DAWLEY; TISSUE=Liver;
RX   MEDLINE [illegible]; PubMed 8335683;
RA   H[illegible] E., Wozniak R.W., Blobel G.;
RT   "An integral membrane protein of the pore membrane domain of the
RT   nuclear envelope contains a nucleoporin-like region.";
RL   J. Cell Biol. 122:513-521(1993).
CC   -!- FUNCTION: ESSENTIAL COMPONENT OF THE NUCLEAR PORE COMPLEX. THE
CC       REPEAT-CONTAINING DOMAIN MAY BE INVOLVED IN ANCHORING COMPONENTS
CC       OF THE PORE COMPLEX TO THE PORE MEMBRANE.
CC   -!- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. NUCLEAR PORE
CC       MEMBRANE.
CC   -!- DOMAIN: CONTAINS X-F-X-F-G REPEATS.
CC   -!- SIMILARITY: THE REPEAT REGION COMPOSED OF PENTAPEPTIDE REPEATS
CC       SEPARATED BY SER/THR-RICH DOMAINS IS SIMILAR TO THAT OF YEAST
CC       NUP1, NSP1 AND MAMMALIAN P62 AND NUP153.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://[illegible]
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; [illegible]; AA79725.1; -.
DR   EMBL; [illegible]; AA79726.1; -.
KW   Nuclear protein; Transport; Transmembrane; Repeat.
FT   DOMAIN        1     56       CISTERNAL SIDE (POTENTIAL).
FT   TRANSMEM     57     77       POTENTIAL.
FT   DOMAIN       78   1199       PORE SIDE (POTENTIAL).
FT   DOMAIN        4      8       POLY ALA.
FT   DOMAIN       53     58       POLY ALA.
FT   DOMAIN      438    441       POLY SER.
FT   DOMAIN      497    500       POLY PRO.
FT   DOMAIN      684    687       POLY SER.
FT   DOMAIN     1023   1026       POLY ALA.
FT   DOMAIN     1033   1038       POLY SER.
FT   DOMAIN     1058   1063       POLY GLY.
SQ   SEQUENCE   1199 AA;  120784 MW;  61C445[illegible] CRC64;

  Query Match             60.0%;  Score [illegible];  DB [illegible];  Length [illegible];
  Best Local Similarity   58.3%;  Pred. No. [illegible];
  Matches   [illegible];  Conservative    2;  Mismatches   [illegible];  Indels   [illegible];  Gaps   [illegible];

Qy          [illegible]

Db          [illegible]


RESULT [illegible]
N[illegible]_HUMAN
ID   N[illegible]_HUMAN     STANDARD;      PRT;  1229 AA.
AC   Q96YN[illegible]; [illegible]; Q9Y4S7;
DT   [illegible] (Rel. 40, Created)
DT   [illegible] (Rel. 40, Last sequence update)
DT   [illegible] (Rel. 40, Last annotation update)
DE   Nuclear envelope pore membrane protein POM 121 (Pore membrane protein
DE   of 121 kDa) (P145).
GN   NUP121 OR KIAA0618.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Cordes M., Bauer C., Holmes A.;
RL   Submitted (NOV-1998) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE OF 243-1229 FROM N.A.
RC   TISSUE=Brain;
RX   MEDLINE 98403880; PubMed 9734811;
RA   Ishikawa K., I., Nagase T., Suyama M., Miyajima N., Tanaka A.,
RA   Kotani H., Nomura N., Ohara O.;
RT   "Prediction of the coding sequences of unidentified human genes. X.
RT   The complete sequences of 100 new cDNA clones from brain which can
RT   code for large proteins in vitro.";
RL   DNA Res. 5:169-176(1998).
RN   [3]
RP   SEQUENCE OF 1130-1229 FROM N.A.
RC   TISSUE=Uterus;
RA   Koehrer K., Beyer A., Mewes H.-W., Gassenhuber J., Wiemann S.;
RL   Submitted (JAN-1999) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: ESSENTIAL COMPONENT OF THE NUCLEAR PORE COMPLEX. THE
CC       REPEAT-CONTAINING DOMAIN MAY BE INVOLVED IN ANCHORING COMPONENTS
CC       OF THE PORE COMPLEX TO THE PORE MEMBRANE (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. NUCLEAR PORE
CC       MEMBRANE (BY SIMILARITY).
CC   -!- DOMAIN: CONTAINS X-F-X-F-G REPEATS.
CC   -!- SIMILARITY: THE REPEAT REGION COMPOSED OF PENTAPEPTIDE REPEATS
CC       SEPARATED BY SER/THR-RICH DOMAINS IS SIMILAR TO THAT OF YEAST
CC       NUP1, NSP1 AND MAMMALIAN P62 AND NUP153.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; AC006014; AAD28064.1; -.
DR   EMBL; AB014518; BAA31593.1; -.
DR   EMBL; AL080109; CAB45714.1; -.
KW   Nuclear protein; Transport; Transmembrane; Repeat.
FT   DOMAIN        1     40       CISTERNAL SIDE (POTENTIAL).
FT   TRANSMEM     41     61       POTENTIAL.
FT   DOMAIN       62   1229       PORE SIDE (POTENTIAL).
FT   DOMAIN        4     10       POLY ALA.
FT   DOMAIN       51     56       POLY ALA.
FT   DOMAIN      294    299       POLY-LYS.
FT   DOMAIN      441    444       POLY SER.
FT   DOMAIN      499    502       POLY PRO.
FT   DOMAIN      733    736       POLY-SER.
FT   DOMAIN      819    826       POLY SER.
FT   DOMAIN      869    875       POLY THR.
FT   DOMAIN     1061   1067       POLY SER.
FT   CONFLICT    297    297       K -> E (IN REF. 2).
FT   CONFLICT    336    336       S -> N (IN REF. 2).
FT   CONFLICT    379    379       T -> A (IN REF. 2).
FT   CONFLICT    451    451       K -> R (IN REF. 2).
FT   CONFLICT    545    545       T -> I (IN REF. 2).
FT   CONFLICT    559    559       T -> P (IN REF. 2).
FT   CONFLICT    702    702       S -> P (IN REF. 2).
FT   CONFLICT    881    881       P -> H (IN REF. 2).
FT   CONFLICT    965    965       T -> A (IN REF. 2).
FT   CONFLICT    982    984       PAT -> A (IN REF. 2).
FT   CONFLICT    991    996       ASIIKI -> PSMIKV (IN REF. 2).
FT   CONFLICT   1000   1000       H -> Y (IN REF. 2).
FT   CONFLICT   1006   1008       QPI -> HPT (IN REF. 2).
FT   CONFLICT   1044   1044       MISSING (IN REF. 2).
FT   CONFLICT   1102   1102       A -> T (IN REF. 2).
```

```
FT   CONFLICT    1124   1124       I -> A (IN REF. 2).
FT   CONFLICT    1146   1146       S   -> G (IN REF. 1).
FT   CONFLICT    1165   1165       L   -> Q (IN REF. 1).
FT   CONFLICT    1197   1197       A   -> G (IN REF. 1).
FT   CONFLICT    1199   1229       SAAPSPESTGAGSKTPGARQRTQAPEQHPPES   -> NTEAHG
FT                                 QEHSPREGPNNLSEREELLPAVEG*TP**GNSPPPPKP
FT                                 (IN REF. 2).
FT   CONFLICT    1234   1234       P   -> L (IN REF. 2).
SQ   SEQUENCE   1239 AA;  135097 MW;  5196551128F8D9D8 CRC64;

  Query Match             98.8%;  Score 89;  DB 1;  Length 1239;
  Best Local Similarity   58.3%;  Pred. No. 18;
  Matches   7;  Conservative    2;  Mismatches    3;  Indels    3;  Gaps    1;

Qy       1 PTPTNSPTLASSE 13
           ||||:|||  ||||
Db     990 PSPSSSPTPGNS 927


RESULT   8
YHB6_ARCFU
ID   YHB6_ARCFU     STANDARD;      PRT:   177 AA.
AC   O28488;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Hypothetical protein AF1786.
GN   AF1786;
OS   Archaeoglobus fulgidus.
OC   Archaea; Euryarchaeota; Archaeoglobales; Archaeoglobaceae;
OC   Archaeoglobus.
OX   NCBI_TaxID=2234;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=VC16 / DSM 4304 / JCM 9629;
RX   MEDLINE=98049344; PubMed=9389475;
RA   Klenk H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,
RA   Ketchum K.A., Dodson R.J., Gwinn M., Hickey E.K., Peterson J.D.,
RA   Richardson D.L., Kerlavage A.R., Graham D.E., Kyrpides N.C.,
RA   Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,
RA   Kirkness E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,
RA   Peterson S., Reich C.I., McNeil L.K., Badger J.H., Glodek A., Zhou L.,
RA   Overbeek R., Gocayne J.D., Weidman J.F., McDonald L., Utterback T.,
RA   Cotton M.D., Spriggs T., Artiach P., Kaine B.P., Sykes S.M.,
RA   Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,
RA   Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
RA   Venter J.C.;
RT   "The complete genome sequence of the hyperthermophilic, sulphate-
RT   reducing archaeon Archaeoglobus fulgidus.";
RL   Nature 390:364-370(1997).
CC   -!- COFACTOR: FMN (By similarity).
CC   -!- SIMILARITY: BELONGS TO THE FLAVODOXIN FAMILY.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; AE000979; AAB89461.1; -.
DR   PIR; A69346; A69346.
DR   InterPro; IPR002563; Flavin_Reduct.
DR   Pfam; PF01613; Flavin_Reduct; 1.
KW   Hypothetical protein; Flavoprotein; FMN; Complete proteome.
SQ   SEQUENCE   177 AA;  19830 MW;  8E48DAF17A480CE3 CRC64;

  Query Match             56.9%;  Score 47;  DB 1;  Length 177;
  Best Local Similarity   58.3%;  Pred. No. 3.9;
  Matches    7;  Conservative    3;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       2 TPGRSPTLASSE 13
           :|||:||  |||
Db     162 SPGKNPTTNSSE 173


RESULT   9
CDH1_HUMAN
ID   CDH1_HUMAN     STANDARD;      PRT:   838 AA.
AC   Q9Y5F3;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
DE   Protocadherin beta 1 precursor (PCDHB1etal).
GN   PCDHB1.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=99308636; PubMed=10380929;
RA   Wu Q., Maniatis T.;
RT   "A striking organization of a large family of human neural cadherin-
RT   like cell adhesion genes.";
RL   Cell 97:779-790(1999).
CC   -!- FUNCTION: POTENTIAL CALCIUM DEPENDENT CELL ADHESION PROTEIN.
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein (by similarity).
CC   -!- SIMILARITY: CONTAINS 6 CADHERIN DOMAINS.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; AF152488; AAD48749.1; -.
DR   HSSP; P15116; 1NCJ.
DR   MIM; 606327; -.
DR   MIM; 604967; -.
DR   InterPro; IPR002126; Cadherin.
DR   Pfam; PF00028; cadherin; 5.
DR   PRINTS; PR00205; CADHERIN.
DR   SMART; SM00112; CA; 6.
DR   PROSITE; PS00232; CADHERIN_1; 5.
DR   PROSITE; PS50268; CADHERIN_2; 6.
KW   Calcium binding; Cell adhesion; Glycoprotein; Signal; Repeat;
KW   Transmembrane; Multigene family.
FT   SIGNAL        1     28       POTENTIAL.
FT   CHAIN        29    838       PROTOCADHERIN BETA 1.
FT   DOMAIN       29    690       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    691    712       POTENTIAL.
FT   DOMAIN      713    838       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN       45    133       CADHERIN 1.
FT   DOMAIN      138    242       CADHERIN 2.
FT   DOMAIN      243    347       CADHERIN 3.
FT   DOMAIN      348    452       CADHERIN 4.
FT   DOMAIN      457    562       CADHERIN 5.
FT   DOMAIN      577    672       CADHERIN 6.
FT   CARBOHYD    169    169       N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD    209    209       N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD    257    257       N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD    419    419       N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD    568    568       N-LINKED (GLCNAC...) (POTENTIAL).
SQ   SEQUENCE   838 AA;  96456 MW;  4EBF7FF4E6A7D6D4 CRC64;

  Query Match             56.9%;  Score 47;  DB 1;  Length 838;
  Best Local Similarity   54.5%;  Pred. No. 27;
  Matches    6;  Conservative    3;  Mismatches    2;  Indels    0;  Gaps    0;
```

Qy 1 FIFGRSFIL 9
Db 308 FIYGFGWIL 308

RESULT 12
OGP_PAPAN
ID OGP_PAPAN STANDARD; PRT; 623 AA.
AC P36718;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Oviduct-specific glycoprotein precursor (oviductal glycoprotein)
DE (oviductin) (Estrogen dependent oviduct protein).
GN OVGP1 OR OGP.
OS Papio anubis (olive baboon).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
OC Cercopithecinae; Papio.
OX NCBI_TaxID=9555;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Oviduct;
RX MEDLINE=98244335; PubMed=9584944;
RA Verhage H.G., Farlentas A.J., Mavrogiannis P.A., Fazleabas A.T.,
RA Donnelly K.M., Arias E.B., Jaffe R.C.;
RT "The baboon oviduct: characteristics of an oestradiol dependent
RT oviduct-specific glycoprotein.";
RL Hum. Reprod. Update 3:541-552(1997).
RN [2]
RP PRELIMINARY PARTIAL SEQUENCE FROM N.A.
RC TISSUE=Oviduct;

OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RA Nhan M.;
RL Submitted (DEC-1995) to the EMBL/GenBank/DDBJ databases.
CC -!- CATALYTIC ACTIVITY: Hydrolysis of the 1,4-beta-linkages of N-
CC acetyl-D-glucosamine polymers of chitin.
CC -!- SIMILARITY: BELONGS TO CHITINASE CLASS II (FAMILY 18 OF GLYCOSYL
CC HYDROLASES).
CC ---------- ----------
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC ---------- ----------
DR EMBL; U42835; AAA83586.1; -.
DR WormPep; C04F6.3; CE03923.
DR InterPro; IPR002557; Chitin_binding.
DR InterPro; IPR001579; Chitinase_2.
DR InterPro; IPR001223; Glyco_hydro_18.
DR Pfam; PF01607; Chitin_bind_2; 2.
DR Pfam; PF00704; Glyco_hydro_18; 1.
DR SMART; SM00494; ChtBD2; 2.
DR PROSITE; PS01095; CHITINASE_18; 1.
KW Hypothetical protein; Hydrolase; Glycosidase; Chitin degradation.
FT ACT_SITE 179 179 PROTON DONOR (BY SIMILARITY).
FT DOMAIN 435 478 CHB_HEX.
SQ SEQUENCE 617 AA; 66857 MW; DDA1D2AAA70E941A CRC64;

Query Match 55.4%; Score 86; DB 1; Length 617;
Best Local Similarity 66.2%; Pred. No. 8;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FIFGRSFIL 9
Db 308 FIYGFGWIL 308

RESULT 13
CHI1_CAEEL
ID CHI1_CAEEL STANDARD; PRT; 617 AA.
AC Q21134;
DT 01-NOV-1996 (Rel. 34, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Putative endochitinase (EC 3.2.1.14).
GN C04F6.3.

```
RX   MEDLINE 91367180; PubMed 1716445;
RA   Doherty K.M., Fazleabas A.T., Verhage H.G., Mayfield J.A.;
RA   Jaffe R.C.;
RT   "Cloning of a recombinant complementary DNA to a baboon (Papio
RT   anubis) estrafol-dependent oviduct specific glycoprotein.";
RL   Mol. Endocrinol. 5:356-364(1991).
CC   -!- FUNCTION: BINDS TO OOCYTE ZONA PELLUCIDA IN VIVO, MAY PLAY A ROLE
CC       IN THE FERTILIZATION PROCESS AND/OR EARLY EMBRYONIC DEVELOPMENT.
CC   -!- SUBCELLULAR LOCATION: Secretory granules.
CC   -!- TISSUE SPECIFICITY: OVIDUCT.
CC   -!- DEVELOPMENTAL STAGE: AT THE TIME OF OVULATION.
CC   -!- SIMILARITY: BELONGS TO FAMILY 18 OF GLYCOSYL HYDROLASES.
CC   -----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----------------------------------------------------------------------
DR   EMBL; M59663; AAB39765.1; -.
DR   PIR; A57654; A57654.
DR   InterPro; IPR001579; Chitinase_2.
DR   InterPro; IPR001223; Glyco_hydro_18.
DR   Pfam; PF00704; Glyco_hydro_18; 1.
DR   PROSITE; PS01095; CHITINASE_18; FALSE_NEG.
KW   Glycoprotein; Fertilization; Signal.
FT   SIGNAL        1     21       BY SIMILARITY.
FT   CHAIN        22    628       OVIDUCT SPECIFIC GLYCOPROTEIN.
FT   CARBOHYD     432   432       N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD     441   441       N-LINKED (GLCNAC...) (POTENTIAL).
SQ   SEQUENCE   628 AA;  70251 MW;  9B21CE46IFFF12CM CRC64;


  Query Match             55.4%;  Score 46;  DB 1;  Length 628;
  Best Local Similarity   66.7%;  Pred. No. 21;
  Matches    6;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 PTFGRSFTL 9
           ||:||:| |
Db     261 PIYGRTFRL 269


RESULT  13
OGP_HUMAN
ID   OGP_HUMAN      STANDARD;      PRT:   628 AA.
AC   Q12889; Q15841;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   15-SEP-2001 (Rel. 40, Last annotation update)
DE   Oviduct specific glycoprotein precursor (oviductal glycoprotein)
DE   (oviductin) (Estrogen-dependent oviduct protein).
GN   OVGP1 OR OGP.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Oviduct;
RX   MEDLINE=91319256; PubMed=2019450;
RA   Arias E.B., Verhage H.G., Jaffe R.C.;
RT   "Complementary deoxyribonucleic acid cloning and molecular
RT   characterization of an estrogen-dependent human oviductal
RT   glycoprotein.";
RL   Biol. Reprod. 45:685-694(1991).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Jaffe R.C.;
RL   Submitted (JUN-1997) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: BINDS TO OOCYTE ZONA PELLUCIDA IN VIVO, MAY PLAY A ROLE
CC       IN THE FERTILIZATION PROCESS AND/OR EARLY EMBRYONIC DEVELOPMENT.
CC   -!- SUBCELLULAR LOCATION: Secretory granules.
CC   -!- TISSUE SPECIFICITY: OVIDUCT.
CC   -!- SIMILARITY: BELONGS TO FAMILY 18 OF GLYCOSYL HYDROLASES.
CC   -----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----------------------------------------------------------------------
DR   EMBL; U09550; AAA86946.1; -.
DR   EMBL; U58010; AAB04126.1; -.
DR   EMBL; U58001; AAB04126.1; JOINED.
DR   EMBL; U58002; AAB04126.1; JOINED.
DR   EMBL; U58003; AAB04126.1; JOINED.
DR   EMBL; U58004; AAB04126.1; JOINED.
DR   EMBL; U58005; AAB04126.1; JOINED.
DR   EMBL; U58006; AAB04126.1; JOINED.
DR   EMBL; U58007; AAB04126.1; JOINED.
DR   EMBL; U58008; AAB04126.1; JOINED.
DR   EMBL; U58009; AAB04126.1; JOINED.
DR   MIM; 603578; -.
DR   InterPro; IPR001579; Chitinase_2.
DR   InterPro; IPR001223; Glyco_hydro_18.
DR   Pfam; PF00704; Glyco_hydro_18; 1.
DR   PROSITE; PS01095; CHITINASE_18; FALSE_NEG.
KW   Glycoprotein; Fertilization; Signal.
FT   SIGNAL        1     21       BY SIMILARITY.
FT   CHAIN        22    628       OVIDUCT SPECIFIC GLYCOPROTEIN.
FT   CARBOHYD     432   432       N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD     441   441       N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD     580   580       N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD     596   596       N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD     648   648       N-LINKED (GLCNAC...) (POTENTIAL).
FT   CONFLICT     422   422       M -> I (IN REF. 2).
FT   CONFLICT     511   511       S -> P (IN REF. 2).
FT   CONFLICT     514   514       Y -> H (IN REF. 2).
FT   CONFLICT     676   676       E -> Q (IN REF. 2).
SQ   SEQUENCE   628 AA;  70251 MW;  9B21CE46IFFF12CM CRC64;


  Query Match             55.4%;  Score 46;  DB 1;  Length 628;
  Best Local Similarity   66.7%;  Pred. No. 21;
  Matches    6;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 PTFGRSFTL 9
           ||:||:| |
Db     261 PIYGRTFRL 269


RESULT  14
OTOF_HUMAN
ID   OTOF_HUMAN     STANDARD;      PRT:  1997 AA.
AC   Q9HC10; Q9HC09; Q9Y650; Q9HC08;
DT   01-MAR-2002 (Rel. 41, Created)
DT   01-MAR-2002 (Rel. 41, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
DE   Otoferlin (Fer-1-like protein 2).
GN   OTOF OR FER1L2.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORMS 1, 2 AND 3), AND ALTERNATIVE SPLICING.
RC   TISSUE=Brain;
RX   MEDLINE=20395831; PubMed=10903124;
RA   Yasunaga S., Grati M., Chardenoux S., Smith T.N., Friedman T.B.,
RA   Lalwani A.K., Wilcox E.R., Petit C.;
RT   "OTOF encodes multiple long and short isoforms: genetic evidence that
RT   the long ones underlie recessive deafness DFNB9.";
```

```
ID   AC   ;  Bad  Seed:  67/501 seq(2000).
XX
AC   ;
AC   Q_[3]N[3] fr M.N.A. (ISOFORM 4).
XX
DT   15-JUL-2001 (Rel. 39, Created)
DT   15-JUL-2001 (Rel. 39, Last sequence update)
DT   15-JUL-2001 (Rel. 39, Last annotation update)
XX
DE   MEMBRANE PROTEIN XP07 PRECURSOR; PubMed Id=10923885;
DE   Yasunaga S., Grati M., Cohen-Salmon M., El-Amraoui A., Mustapha M.,
DE   Salem N., El-Zir E., Loiselet J., Petit C.;
DE   "A mutation in OTOF, encoding otoferlin, a FER-1-like protein, causes
DE   DFNB9, a nonsyndromic form of deafness.";
DE   Nat. Genet. 21:363-369(1999).
DE   -!- FUNCTION: Might be involved in the Ca(2+) triggered synaptic
DE       vesicle-plasma membrane fusion.
DE   -!- SUBCELLULAR LOCATION: Type II membrane protein (By similarity).
DE   -!- ALTERNATIVE PRODUCTS: At least 4 isoforms; 1/short (Shown here),
DE       2/short 2 and 4/short 3, are produced by alternative
DE       splicing.
DE   -!- TISSUE SPECIFICITY: Strongly expressed in brain and inner ear. In
DE       the inner ear, it is mainly expressed in the cochlear IHC and
DE       vestibular type I sensory hair cells. Weakly expressed in eye,
DE       heart, skeletal muscle, liver, kidney, lung and testis.
DE   -!- SIMILARITY: BELONGS TO THE FERLIN FAMILY.
DE   -!- SIMILARITY: CONTAINS 4 C2 DOMAINS.
DE   ---------- ------ ----- - ---- -- -- -- -- -- -- -- -- -- -- -- -- -- -
DE   This SWISS-PROT entry is copyright. It is produced through a collaboration
DE   between the Swiss Institute of Bioinformatics and the EMBL outstation -
DE   the European Bioinformatics Institute. There are no restrictions on its
DE   use by non-profit institutions as long as its content is in no way
DE   modified and this statement is not removed. Usage by and for commercial
DE   entities requires a license agreement (See http://www.isb-sib.ch/announce/
DE   or send an email to license@isb-sib.ch).
DE   ---------- ------ ----- - ---- -- -- -- -- -- -- -- -- -- -- -- -- -- -
DR   EMBL; AF183284; AAG12989.1; -.
DR   EMBL; AF183284; AAG12990.1; -.
DR   HSSP; P04410; 1A25.
DR   MGD; MGI:1891247; Otof.
DR   InterPro; IPR000008; C2.
DR   Pfam; PF00168; C2; 4.
DR   PRINTS; PR00360; C2DOMAIN.
DR   SMART; SM00239; C2; 6.
DR   PROSITE; PS00479; C2_DOMAIN_1; 2.
DR   PROSITE; PS50004; C2_DOMAIN_2; 4.
KW   Transmembrane; Repeat; Alternative splicing.
FT   DOMAIN       1    1963       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM  1964    1984       POTENTIAL.
FT   DOMAIN    1985    1997       EXTRACELLULAR (POTENTIAL).
FT   DOMAIN     241     338       C2 DOMAIN 1.
FT   DOMAIN     403     514       C2 DOMAIN 2.
FT   DOMAIN     947    1052       C2 DOMAIN 3.
FT   DOMAIN    1379    1577       C2 DOMAIN 4.
FT   DOMAIN    1503    1510       POLY-LYS.
FT   DOMAIN    1514    1520       POLY-GLU.
FT   DOMAIN    1906    1994       POLY-LEU.
FT   VARSPLIC     1     747       MISSING (IN ISOFORM 2 AND ISOFORM 3).
FT   VARSPLIC  1445    1454       MISSING (IN ISOFORM 2 AND ISOFORM 4).
FT   VARSPLIC     1     606       MISSING (IN ISOFORM 4).
FT   VARIANT   607     748       MRFQVTDRNYFELLPYLEFKKIGYYKSWWESRFFYLYNANIM
FT                               DHIADKL -> MMIFQLDGDSESSGIMRSLILLLNREEAFG
FT                               EAGEAGIWPSTHIISG (IN ISOFORM 4).
FT   VARIANT   943    1997       SFIWFLNPLKCARYCIRHIYPWCIKIIIIII ILLALFL
FT                               YSVFCYLVKKLIGA -> AFVKFLNECKSIKYLCTKYKWL
FT                               IIKIVLALIGFIMLGLFIYSLFYMVEKLIGA (IN
FT                               ISOFORM 2).
FT   INIT_MET     1    1998       P -> S (IN ELL- 1: AA?).
SQ   SEQUENCE  1997 AA;  226745 MW;  00D1DC8D 2D0C0AAF CRC64;

     Query Match          95.4%;  Score 96;  DB 1;  Length 1997;
     Best Local Similarity  50.0%;  Pred. No. 1.2e+02;
     Matches   27;  Conservative    0;  Mismatches   27;  Indels    0;  Gaps    0;

Qy       1 PIFGRSFILASS 12
           | | :|| : :|
Db    1549 PVFGKSFDIEAS 1560


RESULT  15
OTOF_MOUSE
ID   OTOF_MOUSE     STANDARD;      PRT;  1997 AA.
AC   Q9ESF1; Q9ESF2;
DT   01-MAR-2002 (Rel. 41, Created)
DT   01-MAR-2002 (Rel. 41, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
DE   Otoferlin (Fer-1-like protein 2).
GN   OTOF OR FER1L2.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORM 1 AND 2), AND ALTERNATIVE SPLICING.
RC   TISSUE=Brain, and Cochlea;
RX   MEDLINE=20395851; PubMed=10908124;
RA   Yasunaga S., Grati M., Chardenoux S., Smith T.N., Friedman T.B.,
RA   Lalwani A.K., Wilcox E.R., Petit C.;
RT   "OTOF encodes multiple long and short isoforms: genetic evidence that
RT   the long ones underlie recessive deafness DFNB9.";
RL   Am. J. Hum. Genet. 67:591-600(2000).
RN   [2]
RP   TISSUE SPECIFICITY.
RX   MEDLINE=99206603; PubMed=10092385;
RA   Yasunaga S., Grati M., Cohen-Salmon M., El-Amraoui A., Mustapha M.,
RA   Salem N., El-Zir E., Loiselet J., Petit C.;
RT   "A mutation in OTOF, encoding otoferlin, a FER-1-like protein, causes
RT   DFNB9, a nonsyndromic form of deafness.";
RL   Nat. Genet. 21:363-369(1999).
CC   -!- FUNCTION: Might be involved in the Ca(2+) triggered synaptic
CC       vesicle-plasma membrane fusion.
CC   -!- SUBCELLULAR LOCATION: Type II membrane protein (By similarity).
CC   -!- ALTERNATIVE PRODUCTS: At least 2 isoforms: 1 (shown here) and 2;
CC       are produced by alternative splicing.
CC   -!- TISSUE SPECIFICITY: Strongly expressed in brain and inner ear. In
CC       the inner ear, it is mainly expressed in the cochlear IHC and
CC       vestibular type I sensory hair cells. Weakly expressed in eye,
CC       heart, skeletal muscle, liver, kidney, lung and testis.
CC   -!- SIMILARITY: BELONGS TO THE FERLIN FAMILY.
CC   -!- SIMILARITY: CONTAINS 4 C2 DOMAINS.
CC   ---------- ------ ----- - ---- -- -- -- -- -- -- -- -- -- -- -- -- -- -
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------- ------ ----- - ---- -- -- -- -- -- -- -- -- -- -- -- -- -- -
DR   EMBL; AF183283; AAG12989.1; -.
DR   EMBL; AF183284; AAG12990.1; -.
DR   HSSP; P04410; 1A25.
DR   MGD; MGI:1891247; Otof.
DR   InterPro; IPR000008; C2.
DR   Pfam; PF00168; C2; 4.
DR   PRINTS; PR00360; C2DOMAIN.
DR   SMART; SM00239; C2; 6.
DR   PROSITE; PS00479; C2_DOMAIN_1; 2.
DR   PROSITE; PS50004; C2_DOMAIN_2; 4.
KW   Transmembrane; Repeat; Alternative splicing.
FT   DOMAIN       1    1963       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM  1964    1984       POTENTIAL.
FT   DOMAIN    1985    1997       EXTRACELLULAR (POTENTIAL).
FT   DOMAIN     240     337       C2 DOMAIN 1.
FT   DOMAIN     403     514       C2 DOMAIN 2.
FT   DOMAIN     946    1051       C2 DOMAIN 3.
```

```
FT   DOMAIN      147..   177        /2 DOMAIN 4.
FT   DOMAIN      1505    1513       POLY-LYS.
FT   DOMAIN      1514    1520       POLY-GLU.
FT   DOMAIN      1521    1578       POLY-LEU.
FT   VARSPLIC     169     169       R   > SKGREEEKGGREGEHK (IN ISOFORM 2).
FT   VARSPLIC    1234    1263       MISSING (IN ISOFORM 2).
FT   VARSPLIC    1964    1997       SFIWFLNPLKQAKYFLKHIYHVCLKFIIFCIIILFALFI
FT                                  YSLPGYLAKEIDCA    AFVWFLNGGRSIEYLLCIPVYKWI
FT                                  IIKIVLALLGLLMLALFIYSLFCYMVKKFIGA (IN
FT                                  ISOFORM 2).
SQ   SEQUENCE   1997 AA;  227031 MW;  7B96CA918F79D4D1 CRC64;


  Query Match             55.4%;  Score 46;  DB 1;  Length 1997;
  Best Local Similarity   50.0%;  Pred. No. 1.2e+02;
  Matches    6;  Conservative    3;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 PIPSRSPIIASS 12
           | |||:| : ||
Db    1514 PVPSRSPFPAS 1550


Search completed: August 26, 2002, 06:29:05
Job time: 931 sec
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 26, 2002, 06:28:08 ; Search time 1.5 Seconds
(without alignments)
17.24 Million cell updates/sec

Title: US-09-744-282-5
Perfect score: 65
Sequence: 1 PTEGRSFTLASSE 13

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 95222 seqs, 17299429 residues

Total number of hits satisfying chosen parameters: 95222

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_19:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Prod. N : is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	65	100.0	483	4	Q96H17	Q96h17 homo sapien
2	62	95.4	352	11	Q9WTV1	Q9wtv1 rattus norv
3	62	95.4	381	11	Q99JB4	Q99jb4 mus musculu
4	61	93.8	483	6	Q29411	Q29411 sus scrofa
5	58	89.2	332	6	O18949	O18949 bos taurus
6	57	87.7	387	4	Q9H3V8	Q9h3v8 homo sapien
7	57	87.7	396	11	Q9D7Q1	Q9d7q1 mus musculu
8	57	87.7	466	4	Q14231	Q14231 homo sapien
9	54	83.1	390	4	Q96F97	Q96f97 homo sapien
10	51	78.5	696	5	Q9V2V2	Q9v2v2 drosophila
11	51	78.5	1013	5	Q960M0	Q960m0 drosophila
12	49	75.4	468	5	Q9W2M7	Q9w2m7 drosophila
13	47	72.3	4498	5	Q9W223	Q9w223 drosophila
14	45	69.2	525	5	O44079	O44079 anopheles g
15	43	66.2	558	5	P91731	P91731 lymantria
16	43	66.2	1080	5	Q9BL16	Q9bl16 bombyx mori
17	41	63.1	276	5	Q9V1Z4	Q9v1z4 drosophila
18	41	63.1	442	5	Q9W4V1	Q9w4v1 drosophila
19	41	63.1	574	5	O17411	O17411 aedes aegyp
20	41	63.1	599	5	Q9VFK3	Q9vfk3 drosophila
21	40	61.5	415	16	Q9AA52	Q9aa52 salicornia
22	40	61.5	543	5	Q96GV6	Q96gv6 bombyx mori
23	40	61.5	543	5	Q96GK3	Q96gk3 bombyx mori
24	40	61.5	544	5	Q96Qm4	Q96qm4 bombyx mori
25	40	61.5	552	5	Q96GV44	Q96gv44 spodoptera
26	40	61.5	563	5	Q96GM9	Q96gm9 bombyx mandi
27	40	61.5	565	5	P96710	P96710 bombyx mori
28	40	61.5	1635	5	O17412	O17412 aedes aegyp
29	39	60.0	113	13	Q91153	Q91153 notophthalm
30	39	60.0	280	13	Q91152	Q91152 notophthalm
31	39	60.0	329	4	Q96D10	Q96d10 homo sapien
32	39	60.0	462	5	O15993	O15993 petromyz sp
33	39	60.0	488	5	Q23737	Q23737 chelonus sp
34	39	60.0	497	5	Q25615	Q25615 onchocerca
35	39	60.0	523	4	Q9H9X1	Q9h9x1 homo sapien
36	39	60.0	984	4	Q96HA1	Q96ha1 homo sapien
37	39	60.0	1294	4	Q8Y2N2	Q8y2n2 homo sapien
38	38	58.5	405	5	O97403	O97403 phaedon coc
39	38	58.5	567	10	Q41X71	Q41x71 arabidopsis
40	38	58.5	692	9	Q94M83	Q94m83 streptococc
41	37	56.9	205	17	O29223	O29223 archaeoglob
42	37	56.9	368	5	Q9W2M5	Q9w2m5 drosophila
43	37	56.9	532	2	Q93FK1	Q93fk1 cowdria rum
44	37	56.9	610	5	O74314	O74314 schizosacch
45	37	56.9	641	10	O65642	O65642 arabidopsis

ALIGNMENTS

RESULT 1
Q96H17
ID Q96H17 PRELIMINARY; PRT; 483 AA.
AC Q96H17;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE SIMILAR TO CHITINASE-LIKE 1 (CARTILAGE GLYCOPROTEIN 39).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BRAIN, AND GLIOBLAST-MA WITH ESE AMPLIFICATION;
RA Strausberg R.;
RL Submitted (MAY-2001) to the EMBL/GenBank/DDBJ databases.
DR EMBL; BC008568; AAH08568.1; -.
SQ SEQUENCE 483 AA; 42625 MW; [illegible] CRC64;

Query Match 100.0%; Score 65; DB 4; Length 483;
Best Local Similarity 100.0%; Pred. No. 1.00e-28;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PTEGRSFTLASSE 13
||||||||||||| |
Db 259 PTEGRSFTLASSE 271

RESULT 2
Q9WTV1
ID Q9WTV1 PRELIMINARY; PRT; 352 AA.
AC Q9WTV1;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE GLYCOPROTEIN-39 (FRAGMENT).
OS Rattus norvegicus (Rat).

```
 AT   Eukary; Eu; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 AT   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
  X   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=BRAIN;
RA   Wellington K., Goods A.M.H., Van Eden W.;
RT   "Identification of the rat homologue of human cartilage glycoprotein 39 a
RT   potential autoantigen in arthritis.";
RL   Submitted (APR 1998) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF045610; AAD26610.1; -.
DR   MGD; MGI:1341854; CHI3L1.
DR   InterPro; IPR000677; GH18_Chitinase.
DR   InterPro; IPR001223; Glyco_hydro_18; 1.
DR   Pfam; PF00704; Glyco_hydro_18; 1.
DR   PROSITE; PS51041; GH18_2; 1.
KW   Secreted; Signal.
FT   SIGNAL       1     18
FT   CHAIN       19    382       CHITINASE-3-LIKE PROTEIN 1.
SQ   SEQUENCE   382 AA;  42979 MW;  7B6E6C61CA7B3B47 CRC64;


  Query Match             95.4%;  Score 62;  DB 11;  Length 472;
  Best Local Similarity   92.3%;  Pred. No. 0.00096;
  Matches   12;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 PTFGRSFTLASSE 13
           ||||:|||||||
Db     249 PTFGKSFTLASSE 261


RESULT   4
Q70084
ID   Q70084       PRELIMINARY;      PRT;   381 AA.
AC   Q70084;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   15-DEC-2001 (TrEMBLrel. 18, Last annotation update)
DE   CHEMICAL I CHITINASE 3 LIKE 1 (CARTILAGE GLYCOPROTEIN 39).
GN   Chi3l1.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Hakala B.E.;
RT   Submitted (FEB 2001) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Hakala B.E.;
RT   Submitted (MAR 2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC015411; AAH15411.1; -.
DR   EMBL; BC003280; AAH03280.1; -.
DR   EMBL; BC004734; AAH04734.1; -.
DR   InterPro; IPR001223; Glyco_hydro_18; 1.
DR   Pfam; PF00704; Glyco_hydro_18; 1.
SQ   SEQUENCE   381 AA;  42979 MW;  EF05B8C5AE0D445C CRC64;


  Query Match             95.4%;  Score 62;  DB 11;  Length 381;
  Best Local Similarity   92.3%;  Pred. No. 0.0077;
  Matches   12;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 PTFGRSFTLASSE 13
           ||||:|||||||
Db     249 PTFGKSFTLASSE 261


RESULT   5
Q29411
ID   Q29411       PRELIMINARY;      PRT;   383 AA.
AC   Q29411;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   38 KDA HEPARIN BINDING GLYCOPROTEIN.
OS   Sus scrofa (Pig).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX   NCBI_TaxID=9823;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=SMOOTH MUSCLE;
RX   MEDLINE=95286599; PubMed=7769902;
RA   Shackelton L.M., Mann D.M., Millis A.J.;
RT   "Identification of a 38-kDa heparin-binding glycoprotein (gp38k) in
RT   differentiating vascular smooth muscle cells as a member of a group of
RT   proteins associated with tissue remodeling.";
RL   J. Biol. Chem. 270:13076-13083(1995).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=SMOOTH MUSCLE;
RA   Shackelton L.M., Mann D.M., Millis A.J.;
RL   Submitted (JAN 1997) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U19900; AAA86482.1; -.
DR   EMBL; Z47803; CAA87764.1; -.
DR   InterPro; IPR001223; Glyco_hydro_18.
DR   Pfam; PF00704; Glyco_hydro_18; 1.
KW   Heparin-binding.
SQ   SEQUENCE   383 AA;  42443 MW;  4D1039E4991C8BC4 CRC64;


  Query Match             93.8%;  Score 61;  DB 6;  Length 383;
  Best Local Similarity   92.3%;  Pred. No. 0.0016;
  Matches   12;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 PTFGRSFTLASSE 13
           ||||:|||||||
Db     259 PTFGKSFTLASSK 271


RESULT   6
Q18949
ID   Q18949       PRELIMINARY;      PRT;   332 AA.
AC   Q18949;
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   CHITINASE-LIKE PROTEIN 1 (FRAGMENT).
GN   CLP 1.
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Rocklies A.D., White C.;
RT   "Expression of chitinase like protein 1 (CLP 1) in bovine
RT   chondrocytes.";
RL   Submitted (JUN 1997) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF013174; AAB04364.1; -.
DR   InterPro; IPR001223; Glyco_hydro_18.
DR   Pfam; PF00704; Glyco_hydro_18; 1.
FT   NON_TER       1       1
FT   NON_TER     332     332
SQ   SEQUENCE   332 AA;  37414 MW;  [illegible] CRC64;


  Query Match             89.2%;  Score 58;  DB 6;  Length 332;
  Best Local Similarity   84.6%;  Pred. No. 0.0053;
  Matches   11;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;
```

```
Qy       1 PTPGRSFTLASSE 12
             ||:||||||||
Db     265 PTYGRSFTLASSK 276


RESULT   6
Q8H3V8
ID   Q8H3V8      PRELIMINARY;      PRT:   387 AA.
AC   Q8H3V8;
DT   01-MAR-2002 (TrEMBLrel. 16, Created)
DT   01-MAR-2002 (TrEMBLrel. 16, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   CHITOTRIOSIDASE PRECURSOR.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=96064695; PubMed=7592832;
RA   Boot R.G., Renkema G.H., Verhoek M., Strijland A., Bliek J.,
RA   de Meulemeester T.M., Mannens M.M., Aerts J.M.;
RT   "The human chitotriosidase gene. Nature of inherited enzyme
RT   deficiency.";
RL   J. Biol. Chem. 273:25680-25685(1998).
CC   -!- SIMILARITY: BELONGS TO CHITINASE CLASS II (FAMILY 18 OF GLYCOSYL
CC       HYDROLASES).
DR   EMBL; U62662; AAG10044.1; -.
DR   InterPro; IPR001579; Chitinase_2.
DR   InterPro; IPR001223; Glyco_hydro_18.
DR   Pfam; PF00704; Glyco_hydro_18; 1.
DR   PROSITE; PS01095; CHITINASE_18; 1.
KW   Glycosidase; Hydrolase; Signal.
FT   SIGNAL        1     21       POTENTIAL.
FT   CHAIN        22    387       CHITOTRIOSIDASE.
SQ   SEQUENCE   387 AA;  43133 MW;  03A272B8BC5E0071 CRC64;


  Query Match             87.7%;  Score 57;  DB 4;  Length 387;
  Best Local Similarity   91.7%;  Pred. No. 1.2e-07;
  Matches   11;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 PTPGRSFTLASSE 12
             ||:||||||||
Db     265 PTYGRSFTLASSK 276


RESULT   7
Q9D7Q1
ID   Q9D7Q1      PRELIMINARY;      PRT:   396 AA.
AC   Q9D7Q1;
DT   01-JUN-2000 (TrEMBLrel. 17, Created)
DT   01-JUN-2000 (TrEMBLrel. 17, Last sequence update)
DT   01-JUN-2000 (TrEMBLrel. 17, Last annotation update)
DE   2-oxoglutarate dependent dioxygenase.
GN   2-oxoglutarate.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=BRAIN;
RX   MEDLINE=21085960; PubMed=11217851;
RA   Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA   Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA   Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA   Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA   Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA   Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA   Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA   Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA   Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA   Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA   Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA   Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA   Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA   Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA   Sasaki H., Sato K., Schonbach C., Seya T., Shibata Y., Storch K.F.,
RA   Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA   Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai J., Rossi J.,
RA   Hayashizaki Y.;
RT   "Functional annotation of a full-length mouse cDNA collection.";
RL   Nature 409:685-690(2001).
CC   -!- SIMILARITY: BELONGS TO CHITINASE CLASS II (FAMILY 18 OF GLYCOSYL
CC       HYDROLASES).
DR   EMBL; AK009012; BAB26025.1; -.
DR   MGD; MGI:1919134; 2300002L19Rik.
DR   InterPro; IPR000677; 2S_Globulin.
DR   InterPro; IPR001579; Chitinase_2.
DR   InterPro; IPR001223; Glyco_hydro_18.
DR   Pfam; PF00704; Glyco_hydro_18; 1.
DR   PRINTS; PR00951; 2SGLOBULIN.
DR   PROSITE; PS01095; CHITINASE_18; 1.
KW   Glycosidase; Hydrolase.
SQ   SEQUENCE   396 AA;  43418 MW;  32BB674B6CAF7128 CRC64;


  Query Match             87.7%;  Score 57;  DB 1;  Length 396;
  Best Local Similarity   91.7%;  Pred. No. 1.2e-07;
  Matches   11;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 PTPGRSFTLASSE 12
             ||:||||||||
Db     265 PTYGRSFTLASSK 276


RESULT   8
Q13231
ID   Q13231      PRELIMINARY;      PRT:   466 AA.
AC   Q13231;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   CHITOTRIOSIDASE PRECURSOR.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=96064695; PubMed=7592832;
RA   Boot R.G., Renkema G.H., Strijland A., van Zonneveld A.J., Aerts J.M.;
RT   "Cloning of a cDNA encoding chitotriosidase, a human chitinase
RT   produced by macrophages.";
RL   J. Biol. Chem. 270:26252-26256(1995).
CC   -!- SIMILARITY: BELONGS TO CHITINASE CLASS II (FAMILY 18 OF GLYCOSYL
CC       HYDROLASES).
DR   EMBL; U29615; AAC50246.1; -.
DR   InterPro; IPR001579; Chitinase_2.
DR   InterPro; IPR002557; Chitin_binding.
DR   InterPro; IPR001223; Glyco_hydro_18.
DR   Pfam; PF01607; Chitin_bind_2; 1.
DR   Pfam; PF00704; Glyco_hydro_18; 1.
DR   SMART; SM00494; ChtBD2; 1.
DR   PROSITE; PS01095; CHITINASE_18; 1.
KW   Glycosidase; Hydrolase; Signal.
FT   SIGNAL        1     21       POTENTIAL.
FT   CHAIN        22    466       CHITOTRIOSIDASE.
```

Qy SI, 1N's 1** AA; Sigm MW; B1e12IE**5180..*B*e4;

Query Mat 8% Score 57; DB 1; Length 166;
Best Local Similarity 91.7%; Pred. No. ..12;
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DIIGKFILASS 12
Db 2** DIYGKFILASS 276

RESULT 4
Q9E2
ID Q9E2 PRELIMINARY; PRT; *** AA.
AC Q9E2;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE CHITINASE 2 LIKE 2.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BREAST ANEURYSM WITH EGFP AMPLIFICATION;
RA Zhang J.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AA011460.1; -.
SQ SEQUENCE *** AA; **** MW; ********** CRC64;

Query Match 83.6%; Score 54; DB 1; Length ***;
Best Local Similarity 76.9%; Pred. No. ...*;
Matches 10; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 IIEGKFILASSE 13
Db *** IYGKFILASAE ***

RESULT 5
Q9V7V4
ID Q9V7V4 PRELIMINARY; PRT; 6** AA.
AC Q9V7V4;
DT 01-MAY-2... (TrEMBLrel. 16, Created)
DT 01-MAY-2... (TrEMBLrel. 16, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE ***** PROTEIN.
GN ****.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20...00; PubMed=107...62;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,

RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Lin X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of drosophila melanogaster.";
RL Science 287:2185-2195(2000).
DR EMBL; AE003477; AAF47714.1; -.
DR FlyBase; FBgn0035598; CG1869.
DR InterPro; IPR001579; Chitinase_2.
DR InterPro; IPR001223; Glyco_hydro_18.
DR Pfam; PF00704; Glyco_hydro_18; 2.
DR SMART; SM00636; Glyco_18; DEEN-GN-.
SQ SEQUENCE 696 AA; 77366 MW; 6B775B5B37A19CP6 CRC64;

Query Match 78.5%; Score 51; DB 5; Length 696;
Best Local Similarity 69.2%; Pred. No. 0.26;
Matches 9; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 IIEGKFILASSE 13
Db 668 IIYGKSFILANPD 680

RESULT 11
Q960M0
ID Q960M0 PRELIMINARY; PRT; 1013 AA.
AC Q960M0;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE LD45559p.
GN CG1869.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Y, CN BW SP;
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,
RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,
RA Nunez J., Pacleb J., Paragas V., Park S., Phouanenavong S., Wan K.,
RA Yu C., Lewis S.E., Rubin G.M., Celniker S.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AY051089; AAK93412.1; -.
SQ SEQUENCE 1013 AA; 113285 MW; 26698IAD0685B762 CRC64;

Query Match 78.5%; Score 51; DB 5; Length 1013;
Best Local Similarity 69.2%; Pred. No. 0.39;
Matches 9; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

```
Qy       1 PTFGRSPTLASSE 13
             |:::|||||||:::|
Db     360 PFYCRSPTLAPAE 374


RESULT  12
Q9W2M7
ID   Q9W2M7      PRELIMINARY;      PRT;   460 AA.
AC   Q9W2M7;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   CG9357 PROTEIN.
GN   CG9357.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BERKELEY;
RX   MEDLINE=20196006; PubMed=10731132;
RA   Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA   Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA   George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA   Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA   Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA   Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA   Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA   Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA   Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA   Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA   Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA   Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA   de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA   Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA   Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA   Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA   Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA   Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA   Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA   Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA   Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA   Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA   Liu X., Mattei B., Mcintosh T.C., Mcleod M.P., Mcpherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA   Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA   Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,
RT   "The genome sequence of Drosophila melanogaster";
RL   Science 287:2185-2195(2000).
CC   -!- SIMILARITY: BELONGS TO CHITINASE CLASS II (FAMILY 18 OF GLYCOSYL
CC       HYDROLASES).
DR   EMBL; AE003452; AAF46663.1; -.
DR   HSSP; P07254; 1CTN.
DR   FlyBase; FBgn0034588; CG9357.
DR   InterPro; IPR001579; Chitinase_2.
DR   InterPro; IPR002557; Chitin_binding.
DR   InterPro; IPR001179; FKBP_PPIase.
DR   InterPro; IPR001223; Glyco_hydro_18.
DR   Pfam; PF01607; Chitin_bind_2; 1.
DR   Pfam; PF00704; Glyco_hydro_18; 1.
DR   SMART; SM00494; ChtBD2; 1.
DR   PROSITE; PS01095; CHITINASE_18; 1.
DR   PROSITE; PS00453; FKBP_PPIASE_1; UNKNOWN_1.
KW   Glycosidase; Hydrolase.
SQ   SEQUENCE   460 AA;  51055 MW;  8EF4F126F1DD9D9B CRC64;


  Query Match             75.4%;  Score 49;  DB 5;  Length 460;
  Best Local Similarity   69.2%;  Pred. No. 3.4;
  Matches    9;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 PTFGRSPTLASSE 13
             | :|||||||::|
Db     262 PFYCRSPTLATAE 274


RESULT  13
Q9W2Z3
ID   Q9W2Z3      PRELIMINARY;      PRT;  4498 AA.
AC   Q9W2Z3;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   CG2989 PROTEIN.
GN   CG2989.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BERKELEY;
RX   MEDLINE=20196006; PubMed=10731132;
RA   Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA   Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA   George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA   Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA   Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA   Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA   Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA   Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA   Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA   Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA   Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA   Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA   de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA   Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA   Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA   Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA   Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA   Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA   Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA   Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA   Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA   Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA   Liu X., Mattei B., Mcintosh T.C., Mcleod M.P., Mcpherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA   Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA   Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,
RT   "The genome sequence of Drosophila melanogaster.";
RL   Science 287:2185-2195(2000).
CC   -!- SIMILARITY: BELONGS TO CHITINASE CLASS II (FAMILY 18 OF GLYCOSYL
```

```
CC       GYDROLASES).
DR   EMBL; AF[illegible]; AAF46544.1; -.
DR   PfamBase; PB000343[illegible]; CG2989.
DR   InterPro; IPR001579; Chitinase_2.
DR   InterPro; IPR002557; Chitin_binding.
DR   InterPro; IPR001223; Glyco_hydro_18.
DR   Pfam; PF01607; Chitin_bind_2; 1.
DR   Pfam; PF00704; Glyco_hydro_18; 1.
DR   SMART; SM00494; ChtBD2; 1.
DR   PROSITE; PS01095; CHITINASE_18; 1.
KW   Glycosidase; Hydrolase.
SQ   SEQUENCE   443 AA;  [illegible] MW;  AEB[illegible]AAA[illegible]4C CRC64;


  Query Match             72.4%;  Score 47;  DB 5;  Length 1498;
  Best Local Similarity   61.5%;  Pred. No. 11;
  Matches    8;  Conservative    3;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 PTFGRSFTLAS 11
           : | : |  : |
Db     [illegible] PFYGRGFTLAEE 414


RESULT  14
[illegible]
ID   [illegible]      PRELIMINARY;      PRT;   525 AA.
AC   [illegible];
DT   01-JUN-1998 (TrEMBLrel. 06, Created)
DT   01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   CHITINASE.
GN   AGAP[illegible].
OS   Anopheles gambiae (African malaria mosquito).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea;
OC   Anophelinae.
OX   NCBI_TaxID=7165;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=98360958; PubMed=9660958;
RA   Shen Z., Iatsenko Lorena M.;
RT   "Characterization of a novel gut specific chitinase gene from the
RT   [illegible] malaria vector Anopheles gambiae.";
RL   J. Biol. Chem. 273:28895-28900(1997).
CC   -!- SIMILARITY: BELONGS TO CHITINASE CLASS II (FAMILY 18 OF GLYCOSYL
CC       HYDROLASES).
DR   EMBL; AF[illegible]; AAB87764.1; -.
DR   InterPro; IPR001579; Chitinase_2.
DR   InterPro; IPR002557; Chitin_binding.
DR   InterPro; IPR001223; Glyco_hydro_18.
DR   Pfam; PF01607; Chitin_bind_2; 1.
DR   Pfam; PF00704; Glyco_hydro_18; 1.
DR   SMART; SM00494; ChtBD2; 1.
DR   PROSITE; PS01095; CHITINASE_18; 1.
KW   Glycosidase; Hydrolase.
SQ   SEQUENCE   525 AA;  [illegible] MW;  [illegible]CRC64;


  Query Match             69.2%;  Score 45;  DB 5;  Length 525;
  Best Local Similarity   66.7%;  Pred. No. 27;
  Matches    8;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 PTFGRSFTLAS 11
           .|:|||| |:
Db     [illegible] PLYGRNFLASA 208


RESULT  15
[illegible]
ID   [illegible]      PRELIMINARY;      PRT;   555 AA.
AC   [illegible];
DT   01-MAY-1997 (TrEMBLrel. 04, Created)
```

```
DT   01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   CHITINASE PROTEIN.
OS   Hyphantria cunea (Fall webworm).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
OC   Noctuoidea; Arctiidae; Hyphantria.
OX   NCBI_TaxID=39466;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=98318808; PubMed=9654749;
RA   Kim M.G., Shin S.W., Bae K.S., Kim S.C., Park H.Y.;
RT   "Molecular cloning of chitinase cDNAs from the silkworm, Bombyx mori
RT   and the fall webworm, Hyphantria cunea.";
RL   Insect Biochem. Mol. Biol. 28:163-171(1998).
CC   -!- SIMILARITY: BELONGS TO CHITINASE CLASS II (FAMILY 18 OF GLYCOSYL
CC       HYDROLASES).
DR   EMBL; U86877; AAB47539.1; -.
DR   InterPro; IPR001579; Chitinase_2.
DR   InterPro; IPR002557; Chitin_binding.
DR   InterPro; IPR001223; Glyco_hydro_18.
DR   Pfam; PF01607; Chitin_bind_2; 1.
DR   Pfam; PF00704; Glyco_hydro_18; 1.
DR   SMART; SM00494; ChtBD2; 1.
DR   PROSITE; PS01095; CHITINASE_18; 1.
KW   Glycosidase; Hydrolase.
SQ   SEQUENCE   553 AA;  61968 MW;  5E602B9186AC8FE3 CRC64;


  Query Match             66.2%;  Score 43;  DB 5;  Length 553;
  Best Local Similarity   72.7%;  Pred. No. 7;
  Matches    8;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 PTFGRSFTLAS 11
           |:|||||||:
Db     269 PFYGRSFTLAA 279


Search completed: August 26, 2002, 06:28:08
Job time: 1034 sec
```

GenCore version 4.5
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 26, 2002, 06:13:25 ; Search time 19.28 Seconds
(without alignments)
10,367 Million cell updates/sec

Title: US-09-744-282-5
Perfect score: 65
Sequence: 1 PTFGRSFTLASSE 13

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_032802:*
1: /SIS2/geneseq/ab_if_pro/ab_if_pro.seq:* ...
2: /SIS2/geneseq/ab_if_pro/ab_if_pro.seq:* ...
3: /SIS2/geneseq/ab_if_pro/ab_if_pro.seq:* ...
4: /SIS2/geneseq/ab_if_pro/ab_if_pro.seq:* ...
5: /SIS2/geneseq/ab_if_pro/ab_if_pro.seq:* ...
6: /SIS2/geneseq/ab_if_pro/ab_if_pro.seq:* ...
7: /SIS2/geneseq/ab_if_pro/ab_if_pro.seq:* ...
8: /SIS2/geneseq/ab_if_pro/ab_if_pro.seq:* ...
9: /SIS2/geneseq/ab_if_pro/ab_if_pro.seq:* ...
10: /SIS2/geneseq/ab_if_pro/ab_if_pro.seq:* ...
11: /SIS2/geneseq/ab_if_pro/ab_if_pro.seq:* ...
12: /SIS2/geneseq/ab_if_pro/ab_if_pro.seq:* ...
13: /SIS2/geneseq/ab_if_pro/ab_if_pro.seq:* ...
14: /SIS2/geneseq/ab_if_pro/ab_if_pro.seq:* ...
15: /SIS2/geneseq/ab_if_pro/ab_if_pro.seq:* ...
16: /SIS2/geneseq/ab_if_pro/ab_if_pro.seq:* ...
17: /SIS2/geneseq/ab_if_pro/ab_if_pro.seq:* ...
18: /SIS2/geneseq/ab_if_pro/ab_if_pro.seq:* ...
19: /SIS2/geneseq/ab_if_pro/ab_if_pro.seq:* ...
20: /SIS2/geneseq/ab_if_pro/ab_if_pro.seq:* ...
21: /SIS2/geneseq/ab_if_pro/ab_if_pro.seq:* ...
22: /SIS2/geneseq/ab_if_pro/ab_if_pro.seq:* ...

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	65	100.0	13	17	AAR95075	Peptide conta. hC
2	65	100.0	13	21	AAY69240	Peptide fragment o
3	65	100.0	16	18	AAW45139	Human cartilage gl
4	65	100.0	105	22	AAU02074	Synthetic human ca
5	65	100.0	220	22	AAU02078	Synthetic human mu
6	65	100.0	281	22	AAU02079	Synthetic multi-ta
7	65	100.0	350	22	AAU02076	Synthetic human mu
8	65	100.0	459	22	AAU02077	Synthetic multi-ta
9	61	93.8	383	21	AAB03442	Spark protein Seq
10	58	89.2	388	18	AAW26751	Bovine whey protei
11	57	87.7	473	19	AAW40262	Human chitinase pr
12	57	87.7	473	19	AAW40261	Human chitinase pr
13	57	87.7	473	20	AAY42427	Clone of the C ter
14	57	87.7	473	20	AAY42428	Chitinase amino ac
15	57	87.7	487	18	AAW08585	Human 39 kDa chiti
16	57	87.7	464	18	AAW42447	Human cartilage gl
17	57	87.7	466	18	AAW42448	Human cartilage gl
18	57	87.7	466	18	AAW08584	Human 39 kDa chiti
19	57	87.7	466	19	AAW40259	Human chitinase pr
20	57	87.7	466	19	AAW40260	Human chitinase pr
21	57	87.7	466	20	AAY42425	MC-138 clone of hu
22	57	87.7	466	20	AAY42426	MC-138 clone of hu
23	57	87.7	466	22	AAE00432	Human chitinase pr
24	57	87.7	466	22	AAE00433	Human chitinase pr
25	54	83.1	385	19	AAW53123	Amino acid sequenc
26	54	83.1	385	19	AAW47043	Human cartilage gp
27	54	83.1	416	19	AAW47044	Human cartilage gp
28	54	83.1	421	19	AAW53121	Amino acid sequenc
29	54	83.1	421	19	AAW53122	Amino acid sequenc
30	54	83.1	423	18	AAW25040	Human cartilage gl
31	53	81.5	15	19	AAW61598	Human cartilage gp
32	51	78.5	696	22	ABB58015	Drosophila melanog
33	49	75.4	13	21	AAB03439	Sp38k derived pep.
34	49	75.4	460	22	ABB64466	Drosophila melanog
35	47	72.3	16	18	AAW45140	Human cartilage gl
36	47	72.3	16	19	AAW61599	Human cartilage gp
37	47	72.3	199	22	ABB56452	Human single chain
38	47	72.3	199	22	ABB56456	Human single chain
39	47	72.3	208	22	ABB56475	Human single chain
40	47	72.3	210	22	ABB56442	Human single chain
41	47	72.3	211	22	ABB56441	Human single chain
42	47	72.3	212	22	ABB56455	Human single chain
43	47	72.3	212	22	ABB56439	Human single chain
44	47	72.3	212	22	ABB56440	Human single chain
45	47	72.3	213	22	ABB56457	Human single chain

ALIGNMENTS

RESULT 1
AAR95075
ID AAR95075 standard; peptide; 13 AA.
XX
AC AAR95075;
XX
DT 17-JUL-1996 (first entry)
XX
DE Peptide conta. hC ap 39; autoantigen; rheumatoid arthritis; diagnosis; monitoring; therapy.
XX
KW Human cartilage glycoprotein 39; hC ap 39; autoantigen;
KW inflammation; T cell mediated; destruct.; rheumatoid arthritis; disease model; therapy.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Peptide 1..13
FT /note= "hC ap 39 peptide (1)"
XX
PN WO9613517-A1.
XX
PD 09-MAY-1996.
XX
PF 25-OCT-1995; 95WO-EP04201.
XX
PR 07-APR-1995; 95EP-0200886.
PR 27-OCT-1994; 94EP-0203128.
XX
PA (AKZO) AKZO NOBEL NV.
XX
PI Boots AMBL, Verheijden GFM;
XX
DR WPI; 1996-239447/24.

XX
PT Peptide(s) derived from human cartilage gp 39 autoantigen used
PT in antigen specific T cell tolerance, esp. for treatment of
PT rheumatoid arthritis, also useful in producing arthritic animal
PT model
XX
PN WO9744282-A2.
XX
PD 27-NOV-1997; 28pp; English.
XX
CC The present peptides contain a peptide derived from the human
CC cartilage glycoprotein 39 (HC gp 39) autoantigen. They can be used to
CC induce specific T cell tolerance to HC gp 39 in a patient with a
CC T cell mediated cartilage destruction, esp. rheumatoid arthritis
CC (RA). It may also be used to induce RA in mice, etc. (human disease
CC model), and to diagnose the presence of activated T cells
CC reactive with HC gp 39, i.e. to identify patients suitable for
CC specific treatment.
CC The peptide is given in a dose of 1-1 000 micro g/kg
CC by injection, or 10-50 microg/kg to produce a disease model. The
CC peptide has a specific tolerising effect on autoreactive T cells,
CC but leaves other components of the immune system intact, and
CC therefore should have few side effects.
XX
SQ Sequence 13 AA;

Query Match 100.0%; Score 65; DB 12; Length 13;
Best Local Similarity 100.0%; Pred. No. 2.7e-06;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FIPSRKSFILASSE 13
||||||||||||| |
Db 1 FIPSRKSFILASSE 13

RESULT 5
AAW49139
ID AAW49139 standard; peptide; 16 AA.
XX
AC AAW49139;
XX
DT 28-MAR-1998 (first entry)
XX
DE Human cartilage glycoprotein 39 derived peptide #40.
XX
KW Articular cartilage; immunosuppressive therapy; antigen specific;
KW immunological tolerance; T-cell; human cartilage glycoprotein 39;
KW HC gp 39; rheumatoid arthritis; epitope.
XX
OS Synthetic.
OS Homo sapiens.
XX
PN WO9744058-A1.
XX
PD 30-OCT-1997.
XX
PF 22-APR-1997; 97WO-EP02051.
XX
PR 24-APR-1996; 96EP-0201106.
XX
PA (ALKU) AKZO NOBEL NV.
XX
PI Boots AMH; Verheijden GFM;
XX
DR WPI; 1997-535725/49.
XX
PT Peptide suitable for use in antigen specific immunosuppressive
PT therapy - resembles or mimics epitope present on HC gp 39, so
PT inducing systemic immunological tolerance to rheumatoid arthritis
PT auto-antigen
XX
PS Disclosure; Page 16; 82pp; English.
XX
CC The present sequence represents a peptide which resembles or mimics an
CC epitope present on human cartilage glycoprotein 39 (HC gp 39), an
CC autoantigen in rheumatoid arthritis. The invention relates to peptides
CC consisting of 16-55 amino acid residues comprising at least one of the
CC following 19 sequences: LVCYYISWS; FLCIHIIYS; IIYSFANIS; IKCLLSVAS;
CC FIKSVPPSI; PDCLDAWAI; LYPGPRDPQ; YPIAFISQR; LSFISIMIY; FICIMIYDF;
CC FRGGEDASP; YAVGYMLKL; MLKLGAFAS; LAYYEICDF; LAGAIVHKI; YLKDSGLAG;
CC LAGAMVWAL; VWALDLDDF; or LDLDDPQGS. They can be used medically in
CC antigen-specific immunosuppressive therapy, particularly the treatment
CC of T cell mediated destruction of articular cartilage in autoimmune
CC diseases (e.g. rheumatoid arthritis). They can also be used to detect
CC activated autoreactive T cells in an individual. The peptides have a
CC specific effect on the autoreactive T cells, thus leaving the other

```
CC   components of the immune system intact, unlike the non-specific
CC   suppressive effect of immunosuppressive drugs, and also have fewer
CC   side effects. The peptides are predominantly recognised by autoreactive
CC   T cells from rheumatoid arthritis patients, but rarely by those from
CC   healthy donors.
XX
SQ   Sequence   16 AA;


  Query Match             100.0%;  Score 65;  DB 18;  Length 16;
  Best Local Similarity   100.0%;  Pred. No. 9.6e-06;
  Matches   13;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 PTPGRSFILASSE 13
           |||||||||||||
Db       4 ptpgrsfilasse 16


RESULT   4
AAU02074
ID   AAU02074 standard; Protein; 105 AA.
XX
AC   AAU02074;
XX
DT   07-SEP-2001  (first entry)
XX
DE   Synthetic human target autoantigen shGLP/hA.
XX
KW   shGLP/hA; immunogenic epitope cluster; IEC;
KW   synthetic human chondrocyte glycoprotein 60/cartilage link protein;
KW   interstitial autoimmune disease; multiple sclerosis;
KW   rheumatoid arthritis; insulin-dependent diabetes mellitus;
KW   myasthenia gravis; uveitis; autoimmune hepatitis; thyroiditis; insulitis;
KW   orchitis; idiopathic thrombocytopenic purpura; inflammatory disease;
KW   Crohn's disease; ulcerative colitis.
XX
OS   Synthetic.
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 103
FT                   /label= OTHER
FT                   /note= "Other= STOP"
XX
PN   WO200131037-A2.
XX
PD   03-MAY-2001.
XX
PF   26-OCT-2000; 2000WO-IL00688.
XX
PR   27-OCT-1999;  99IL-0132611.
XX
PA   (YEDA ) YEDA RES & DEV CO LTD.
XX
PI   Ben-Nun A,  Kerlero De Rosbo N,  Sappler GE;
XX
DR   WPI; 2001-300515/31.
DR   N-PSDB; AAS04827.
XX
PT   Novel synthetic human target autoantigen gene useful for treating
PT   autoimmune diseases such as multiple sclerosis, insulin dependent
PT   diabetes mellitus, rheumatoid arthritis, myasthenia gravis, and uveitis
PT   .
XX
PS   Claim 178; Fig 51; 182pp; English.
XX
CC   The sequence represents shGLP/hA (synthetic human chondrocyte
CC   glycoprotein 60/cartilage link protein) containing immunogenic epitope
CC   clusters (IEC) from chondrocyte glycoprotein 60 and cartilage link
CC   protein. The synthetic human target autoantigen genes of the invention
CC   comprise sequences coding for at least 2 IECs of autoantigen(s) related
CC   to a specific autoimmune disease. The synthetic human target autoantigen
CC   genes are useful for treating autoimmune diseases such as multiple
CC   sclerosis, insulin-dependent diabetes mellitus, rheumatoid arthritis,
CC   myasthenia gravis, uveitis, autoimmune hepatitis, thyroiditis, insulitis,
CC   orchitis, idiopathic thrombocytopenic purpura, and inflammatory diseases
CC   (Crohn's disease, ulcerative colitis). The synthetic human target
CC   autoantigen genes are also useful for diagnosis and/or monitoring the
CC   progression of the autoimmune disease.
XX
SQ   Sequence   105 AA;


  Query Match             100.0%;  Score 65;  DB 22;  Length 105;
  Best Local Similarity   100.0%;  Pred. No. 7.4e-05;
  Matches   13;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 PTPGRSFILASSE 13
           |||||||||||||
Db      25 ptpgrsfilasse 37


RESULT   5
AAU02078
ID   AAU02078 standard; Protein; 220 AA.
XX
AC   AAU02078;
XX
DT   07-SEP-2001  (first entry)
XX
DE   Synthetic human multi target autoantigen Y-RAPs.
XX
KW   Y-RAPs; immunogenic epitope cluster; IEC;
KW   autoantigen; shCollagen/hA; shAggrecan/hA; shGLP/hA;
KW   autoimmune disease; multiple sclerosis; rheumatoid arthritis;
KW   insulin-dependent diabetes mellitus; myasthenia gravis; uveitis;
KW   autoimmune hepatitis; thyroiditis; insulitis; orchitis;
KW   idiopathic thrombocytopenic purpura; inflammatory disease;
KW   Crohn's disease; ulcerative colitis; shMultiTAG.
XX
OS   Synthetic.
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 218
FT                   /label= OTHER
FT                   /note= "Other= STOP"
XX
PN   WO200131037-A2.
XX
PD   03-MAY-2001.
XX
PF   26-OCT-2000; 2000WO-IL00688.
XX
PR   27-OCT-1999;  99IL-0132611.
XX
PA   (YEDA ) YEDA RES & DEV CO LTD.
XX
PI   Ben-Nun A,  Kerlero De Rosbo N,  Sappler GE;
XX
DR   WPI; 2001-300515/31.
DR   N-PSDB; AAS04839.
XX
PT   Novel synthetic human target autoantigen gene useful for treating
PT   autoimmune diseases such as multiple sclerosis, insulin dependent
PT   diabetes mellitus, rheumatoid arthritis, myasthenia gravis, and uveitis
PT   .
XX
PS   Claim 185; Fig 59; 182pp; English.
XX
CC   The sequence represents synthetic human multi target autoantigen
CC   (shMultiTAG) Y-RAPs consisting of regions encoding preferred epitopes
CC   from shCollagen/hA, shAggrecan/hA and shGLP/hA. The synthetic human
CC   target autoantigen genes of the invention comprise sequences coding for
CC   at least 2 IECs of autoantigen(s) related to a specific autoimmune
CC   disease. The synthetic human target autoantigen genes are useful for
```

```
CC   Treating autoimmune diseases such as multiple sclerosis, insulin
CC   dependent diabetes mellitus, rheumatoid arthritis, myasthenia gravis,
CC   uveitis, autoimmune hepatitis, thyroiditis, insulitis, orchitis,
CC   idiopathic thrombocytopenic purpura, and inflammatory diseases (Crohn's
CC   disease, ulcerative colitis). The synthetic human target autoantigen
CC   genes are also useful for diagnosis and/or monitoring the progression of
CC   the autoimmune disease.
XX
SQ   Sequence   281 AA;


  Query Match             100.0%;  Score 65;  DB 22;  Length 281;
  Best Local Similarity   100.0%;  Pred. No. 0.00023;
  Matches   13;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        1 TTPGRSFTLASSE 13
            | | | |||||||
Db      181 TTPGRSFTLASSE 193


RESULT   7
AA002076
ID   AA002076 standard; Protein; 450 AA.
XX
AC   AA002076;
XX
DT   07-SEP-2001  (first entry)
XX
DE   Synthetic human multi-target autoantigen Y-RAPa.
XX
KW   Y-RAPa; immunogenic epitope cluster; IEC;
KW   autoantigen; sh/villagin/pA; sh/Airtin/pA; sh/Ili/pA;
KW   autoimmune disease; multiple sclerosis; rheumatoid arthritis;
KW   insulin-dependent diabetes mellitus; myasthenia gravis; uveitis;
KW   autoimmune hepatitis; thyroiditis; insulitis; orchitis;
KW   idiopathic thrombocytopenic purpura; inflammatory disease;
KW   Crohn's disease, ulcerative colitis; shMultiAG.
XX
OS   Synthetic.
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 438
FT                   /label  OTHER
FT                   /note  "Other  STOP"
XX
PN   WO200131037 A2.
XX
PD   03 MAY 2001.
XX
PF   26 OCT 2000; 2000WO IL00688.
XX
PR   27-OCT-1999;  99IL 0132611.
XX
PA   (YEDA ) YEDA RES & DEV CO LTD.
XX
PI   Ben Num A,  Kerlero De Rosbo N,  Sappler GP;
XX
DR   WPI; 2001-300515/31.
DR   N-PSDB; AAS04837.
XX
PI   Novel synthetic human target autoantigen gene useful for treating
PI   autoimmune diseases such as multiple sclerosis, insulin dependent
PI   diabetes mellitus, rheumatoid arthritis, myasthenia gravis, and uveitis
PI
XX
PS   Claim 179; Fig 55; 182pp; English.
XX
CC   The sequence represents synthetic human multi target autoantigen
CC   (shMultiAG); Y RAPa consisting of protected epitope cluster (IEC)
CC   target autoantigen genes of the invention comprise sequences coding for
CC   at least 2 IECs of autoantigen(s) related to a specific autoimmune
CC   disease. The synthetic human target autoantigen genes are useful for
CC   treating autoimmune diseases such as multiple sclerosis, insulin
CC   dependent diabetes mellitus, rheumatoid arthritis, myasthenia gravis,
CC   uveitis, autoimmune hepatitis, thyroiditis, insulitis, orchitis,
CC   idiopathic thrombocytopenic purpura, and inflammatory diseases (Crohn's
CC   disease, ulcerative colitis). The synthetic human target autoantigen
CC   genes are also useful for diagnosis and/or monitoring the progression of
CC   the autoimmune disease.
XX
SQ   Sequence   281 AA;


  Query Match             100.0%;  Score 65;  DB 22;  Length 281;
  Best Local Similarity   100.0%;  Pred. No. 0.00023;
  Matches   13;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        1 TTPGRSFTLASSE 13
            | | | |||||||
Db      181 TTPGRSFTLASSE 193


RESULT   7
AA002076
ID   AA002076 standard; Protein; 450 AA.
XX
AC   AA002076;
XX
DT   07-SEP-2001  (first entry)
XX
DE   Synthetic human multi-target autoantigen Y-RAPa.
XX
KW   Y-RAPa; immunogenic epitope cluster; IEC;
KW   autoantigen; sh/villagin/pA; sh/Airtin/pA; sh/Ili/pA;
KW   autoimmune disease; multiple sclerosis; rheumatoid arthritis;
KW   insulin-dependent diabetes mellitus; myasthenia gravis; uveitis;
KW   autoimmune hepatitis; thyroiditis; insulitis; orchitis;
KW   idiopathic thrombocytopenic purpura; inflammatory disease;
KW   Crohn's disease, ulcerative colitis; shMultiAG.
XX
OS   Synthetic.
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 438
FT                   /label  OTHER
FT                   /note  "Other  STOP"
XX
PN   WO200131037 A2.
XX
PD   03 MAY 2001.
XX
PF   26 OCT 2000; 2000WO IL00688.
XX
PR   27-OCT-1999;  99IL 0132611.
XX
PA   (YEDA ) YEDA RES & DEV CO LTD.
XX
PI   Ben Num A,  Kerlero De Rosbo N,  Sappler GP;
XX
DR   WPI; 2001-300515/31.
DR   N-PSDB; AAS04837.
XX
PI   Novel synthetic human target autoantigen gene useful for treating
PI   autoimmune diseases such as multiple sclerosis, insulin dependent
PI   diabetes mellitus, rheumatoid arthritis, myasthenia gravis, and uveitis
PI
XX
PS   Claim 179; Fig 55; 182pp; English.
XX
CC   The sequence represents synthetic human multi target autoantigen
CC   (shMultiAG); Y RAPa consisting of protected epitope cluster (IEC)
```

CC shAqp1/kA and shCLP/kA. The synthetic human target autoantigen genes
CC of the invention comprise sequences coding for at least 2 IECs of
CC autoantigen(s) related to a specific autoimmune disease. The synthetic
CC human target autoantigen genes are useful for treating or immune
CC diseases such as multiple sclerosis, insulin dependent diabetes
CC mellitus, rheumatoid arthritis, myasthenia gravis, uveitis, autoimmune
CC hepatitis, thyroiditis, insulitis, orchitis, idiopathic thrombocytopenic
CC purpura, and inflammatory diseases (Crohn's disease, ulcerative colitis).
CC The synthetic human target autoantigen genes are also useful for
CC diagnosis and/or monitoring the progression of the autoimmune disease.
XX
SQ Sequence 459 AA;

Query Match 100.0%; Score 65; DB 2; Length 459;
Best Local Similarity 100.0%; Pred. No. 0.00027;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PTFGRSPTLASSE 13
||||||||||||| |
Db 270 ptfqrsttlasse 282

RESULT 8
AAB02077
ID AAB02077 standard; Protein; 459 AA.
XX
AC AAB02077;
XX
DT 07-SEP-2000 (first entry)
XX
DE Synthetic multi-target autoantigen Y-RAPa.
XX
KW Y-RAPa; immunogenic epitope cluster; IEC;
KW autoantigen; shAqp1/kA; shAqp1-like/kA; shCLP/kA; hn-/kA;
KW autoimmune disease; multiple sclerosis; rheumatoid arthritis;
KW insulin dependent diabetes mellitus; myasthenia gravis; uveitis;
KW autoimmune hepatitis; thyroiditis; insulitis; orchitis;
KW idiopathic thrombocytopenic purpura; inflammatory disease;
KW Crohn's disease; ulcerative colitis; shMultiTAG.
XX
OS Synthetic.
OS Homo sapiens.
OS Escherichia coli.
OS Mycobacterium tuberculosis.
XX
FH Key Location/Qualifiers
FT Misc-difference 456
FT /label OTHER
FT /note "other STOP"
FT Misc-difference 457
FT /label OTHER
FT /note "other STOP"
XX
PN WO200063367-A2.
XX
PD 26-OCT-2000.
XX
PF 20-APR-2000; 2000WO-IL000688.
XX
PR 27-APR-1999; 99IL-0132611.
XX
PA (YEDA) YEDA RES & DEV CO LTD.
XX
PI Ben-Nun A, Kerlero De Rosbo N, Sappler GB;
XX
DR WPI; 2000-615723.
DR N-PSDB; AAS04838.
XX
PT Novel synthetic human target autoantigen gene useful for treating
PT autoimmune diseases such as multiple sclerosis, insulin dependent
PT diabetes mellitus, rheumatoid arthritis, myasthenia gravis, and uveitis
PT

XX
PS Claim 182; Fig 57; 182pp; English.
XX
CC The sequence represents synthetic human multi target autoantigen
CC (shMultiTAG) Y-RAPa consisting of regions encoding shAqp1/kA,
CC shAqp1-like/kA, hn-/kA and shCLP/kA. The synthetic human target
CC autoantigen genes of the invention comprise sequences coding for at least
CC 2 IECs of autoantigen(s) related to a specific autoimmune disease. The
CC synthetic human target autoantigen genes are useful for treating
CC autoimmune diseases such as multiple sclerosis, insulin dependent
CC diabetes mellitus, rheumatoid arthritis, myasthenia gravis, uveitis,
CC autoimmune hepatitis, thyroiditis, insulitis, orchitis, idiopathic
CC thrombocytopenic purpura, and inflammatory diseases (Crohn's disease,
CC ulcerative colitis). The synthetic human target autoantigen genes are
CC also useful for diagnosis and/or monitoring the progression of the
CC autoimmune disease.
XX
SQ Sequence 459 AA;

Query Match 100.0%; Score 65; DB 2; Length 459;
Best Local Similarity 100.0%; Pred. No. 0.00027;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PTFGRSPTLASSE 13
||||||||||||| |
Db 270 ptfqrsttlasse 282

RESULT 9
AAB03442
ID AAB03442 standard; Protein; 483 AA.
XX
AC AAB03442;
XX
DT 03-JAN-2001 (first entry)
XX
DE Gp38k protein sequence.
XX
KW Gp38k; chemoattractant; cell migration; wound healing; angiogenesis;
KW cancer; vascular trauma; vascular disease; atherosclerosis; restenosis;
XX
OS Unidentified.
XX
FH Key Location/Qualifiers
FT Peptide 1..24
FT /label signal_peptide
FT Region 30..37
FT /note "peptide antibody"
FT Modified-site 59..61
FT /label glycosylation_site
FT Binding site 68..75
FT /label leucine_zipper
FT Binding site 82..89
FT /label leucine_zipper
FT Active site 131..136
FT /label chitanase_active_site
FT Binding site 143..146
FT /label heparin_binding_site
FT Binding site 147..154
FT /label hyaluronic_acid_binding_site
FT Binding site 262..270
FT /label hyaluronic_acid_binding_site
FT Binding site 278..281
FT /label glycosaminoglycan_binding_site
FT Region 354..357
FT /note "acidic region"
FT Binding-site 368..376
FT /label hyaluronic_acid_binding_site
XX
PN WO200034469-A1.
XX
PD 15-JUN-2000.

XX
PF [illegible] Proc [illegible] 96WO US10460.
XX
PR [illegible] 96US 0663618.
XX
PA [illegible] ; INIV NEW YORK STATE RES FOUND.
XX
PI [illegible] All.
XX
DR WPI; [illegible] 431810/47
XX
PT [illegible] and p98K related peptide capable of alterin cell migration
PT [illegible] for treating atherosclerosis, cancer and stenosis following
PT [illegible] vascular trauma or disease
XX
PS [illegible] Fig 2; 43pp; English.
XX
CC The present sequence is the protein sequence of gp98k. Sp98k, a
CC chemoattractant, is essential for the migration of vascular smooth muscle
CC cells (VSMC). The gene and protein can, therefore, be used to promote
CC wound healing, angiogenesis and vasculogenesis, in the treatment of
CC stenosis following vascular trauma or disease and to treat
CC atherosclerosis, and antisense sequences can be used to treat cancer, as
CC angiogenesis is vital for tumour survival.
XX
SQ Sequence [illegible] AA;

Query Match [illegible] 93.8%; Score 61; DB 21; Length 384;
Best Local Similarity [illegible] 92.9%; Pred. No. 0.0048;
Matches [illegible] Conservative [illegible] Mismatches [illegible] Indels [illegible] Gaps [illegible];

Qy [illegible]
[illegible]
Db 259 [illegible] 271

RESULT [illegible]
AAW40262 [illegible]
ID AAW40262 standard; Protein; 384 AA.
XX
AC AAW40262;
XX
DT [illegible] MAY 1998 (first entry)
XX
DE [illegible] whey protein.
XX
KW [illegible] whey protein; human cartilage glycoprotein 39; HC gp 39;
KW [illegible] autoimmune disease; rheumatoid arthritis;
KW [illegible] arthritogenic protein; immunotherapy; therapy.
XX
OS Bos [illegible]
XX
PN WO9747752-A1.
XX
PD [illegible] DEC 1997.
XX
PF [illegible] JUN 1997; 97WO US10460.
XX
PR [illegible] JUN 1996; 96US 0663618.
XX
PA [illegible] ; AKZO NOBEL NV.
XX
PI [illegible] AMB; [illegible] BS; Verbeijden; GEM;
XX
DR WPI; [illegible] 06-6/49.
DR N-PSDB; AAL04[illegible].
XX
PT [illegible] protein homologous to human cartilage glycoprotein 39
PT [illegible] induce arthritis and provide antigen specific treatment of articular
PT [illegible] cartilage destruction in autoimmune diseases eg. rheumatoid
PT [illegible] arthritis
XX
PS Claim 11; Page 23-25; 85pp; English.
XX
CC This bovine 39 kDa whey protein sequence was deduced from cDNA
CC clones (see AAT99452) isolated from a cattle mammary gland cDNA
CC library using human cartilage glycoprotein 39 (HC gp 39) cDNA as
CC probe. Bovine whey protein was shown to be arthritogenic, inducing
CC arthritis in animals in the same way as described for HC gp 39 (see
CC AAW26750). Arthritogenic proteins such as bovine whey protein, and
CC which display at least 50% amino acid homology to HC gp 39, are
CC very suitable for inducing systemic tolerance of the immune system
CC to homologous autoantigens and can be used to delay and/or suppress
CC arthritic development in mammals. They can induce specific T cell
CC tolerance to HC gp 39 in patients suffering from T cell mediated
CC cartilage destruction, such as rheumatoid arthritis. Arthritogenic
CC proteins are also suitable to induce arthritis in animals,
CC preferably mice, e.g. for use in drug screening.
XX
SQ Sequence 383 AA;

Query Match 89.2%; Score 58; DB 18; Length 383;
Best Local Similarity 84.6%; Pred. No. 0.0067;
Matches 11; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PIPGRSFTLASSE 13
[illegible]
Db 259 ptpqrsytlassk 271

RESULT 11
AAW40262
ID AAW40262 standard; Protein; 373 AA.
XX
AC AAW40262;
XX
DT 15-JUN-1998 (first entry)
XX
DE Human chitinase protein analogue.
XX
KW Chitinase; human; fungal infection; immunogen; diagnosis; treatment;
KW Gaucher's disease; transgenic; detection; hybridisation; antifungal;
KW rheumatoid arthritis; overexpression; extracellular matrix.
XX
OS Homo sapiens.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Misc-difference 370
FT /label= P370S
FT /note= "Wild-type Pro is replaced by Ser"
XX
PN WO9747752-A1.
XX
PD 18-DEC-1997.
XX
PF 16-JUN-1997; 97WO US10460.
XX
PR 14-JUN-1996; 96US 0663618.
XX
PA (IOOS-) IOOS CORP.
XX
PI Gray PW;
XX
DR WPI; 1998-052816/05.
XX
PT Nucleic acids encoding human chitinase - useful as antifungal
PT agents, especially in combination with other antifungals
XX
PS Claim 29; Page 49-50; 63pp; English.
XX
CC This sequence encodes a fragment of a novel human chitinase protein
CC analogue in which a proline at position 370 of the wild-type protein
CC (see AAW40261) is substituted with a serine and the C-terminal 22

CC residues of the mature protein are deleted. These residues are not
CC critical to chitinase enzymatic activity. Chitinases are useful for
CC treating or preventing fungal infection and as immunogens for generating
CC antibodies which are used to purify, detect and quantify chitinases, e.g.
CC for diagnosis of Gaucher's disease. The nucleic acid sequence of the
CC chitinase is also useful as a probe to identify and isolate genomic DNA
CC encoding chitinases or similar proteins, or cells expressing them or to
CC generate transgenic ('knockout') rodents. It can also be used in
CC hybridisation assays and to detect genetic alterations in the chitinase
CC gene related to disease. Agents that inhibit this protein may be useful
CC in treatment of Gaucher's disease and rheumatoid arthritis, where
CC overexpression of the protein can damage the extracellular matrix.
CC Chitinase also improves the activity of other antifungal agents and may
CC allow a reduction in the dose of such agents, and thus of their side
CC effects.
XX
SQ Sequence 373 AA;

Query Match 87.7%; Score 57; DB 19; Length 373;
Best Local Similarity 91.7%; Prod. No. 0.01;
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PTPGRSFTLASS 12
|:|:|||||||||
Db 244 ptyqrsftlass 255

RESULT 12
AAW40261
ID AAW40261 standard; Protein; 373 AA.
XX
AC AAW40261;
XX
DT 15-JUN-1998 (first entry)
XX
DE Human chitinase protein fragment.
XX
KW Chitinase; human; fungal infection; immunogen; diagnosis; treatment;
KW Gaucher's disease; transgenic; detection; hybridisation; antifungal;
KW rheumatoid arthritis; overexpression; extracellular matrix.
XX
OS Homo sapiens.
XX
PN WO9746252-A1.
XX
PD 18-DEC-1997.
XX
PF 16-JUN-1997; 97WO-US10460.
XX
PR 14-JUN-1996; 96US-0663618.
XX
PA (ICOS-) ICOS CORP.
XX
PI Gray PW;
XX
DR WPI; 1998-052316/05.
DR N-PSDB; AAV22435.
XX
PT Nucleic acids encoding human chitinase useful as antifungal
PT agents, especially in combination with other antifungals
XX
PS Claim 24; Page 48-49; 63pp; English.
XX
CC This sequence encodes a fragment of a novel human chitinase protein
CC which lacks the C terminal 72 residues of the mature protein. These
CC residues are not critical to chitinase enzymatic activity. Chitinases are
CC useful for treating or preventing fungal infection and as immunogens for
CC generating antibodies which are used to purify, detect and quantify
CC chitinases, e.g. for diagnosis of Gaucher's disease. The nucleic acid
CC sequence of the chitinase is also useful as a probe to identify and
CC isolate genomic DNA encoding chitinases or similar proteins, or cells
CC expressing them or to generate transgenic ('knockout') rodents. It can
CC also be used in hybridisation assays and to detect genetic alterations
CC in the chitinase gene related to disease. Agents that inhibit this
CC protein may be useful in treatment of Gaucher's disease and rheumatoid
CC arthritis, where overexpression of the protein can damage the
CC extracellular matrix. Chitinase also improves the activity of other
CC antifungal agents and may allow a reduction in the dose of such agents,
CC and thus of their side effects.
XX
SQ Sequence 373 AA;

Query Match 87.7%; Score 57; DB 19; Length 373;
Best Local Similarity 91.7%; Prod. No. 0.01;
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PTPGRSFTLASS 12
|:|:|||||||||
Db 244 ptyqrsftlass 255

RESULT 13
AAY42427
ID AAY42427 standard; Protein; 373 AA.
XX
AC AAY42427;
XX
DT 10-DEC-1999 (first entry)
XX
DE Clone of the C terminal fragment of human chitinase
XX
KW chitin; fungal infection; immunocompromised; AIDS; chemotherapy;
KW organ transplant; parasite; chitin binding; allelic vector;
KW truncated protein; bacterial expression.
XX
OS Synthetic.
OS Homo sapiens.
XX
PN WO9946390-A1.
XX
PD 16-SEP-1999.
XX
PF 12-MAR-1999; 99WO-US05343.
XX
PR 12-MAR-1998; 98US-0039198.
XX
PA (ICOS-) ICOS CORP.
XX
PI Gray PW, Tjoelker LW;
XX
DR WPI; 1999-551417/46.
XX
PT Novel chitin-binding fragments of human chitinase used to treat fungal
PT infections in animals
XX
PS Example 5; Page 66-67; 83pp; English.
XX
CC This is the amino acid sequence of the C-terminal fragment of human
CC chitinase, which can be expressed in bacterial cells. The fragment was
CC prepared by amplifying the MG-218 plasmid with primers 218.1 (AAZ21855)
CC and T-END (AAZ21856).
CC Chitinase fragments can be used to screen for proteins or other
CC molecules that specifically bind to the chitin binding domain of human
CC chitinase or that modulate its activity. These components are useful in
CC immunization, as well as for purifying chitinase, as well as for
CC detection and quantification of chitinase. Polynucleotide fragments of
CC the invention are useful as a source of probes and primers, and to
CC express the proteins recombinantly. The chitinase fragments, when
CC conjugated to antifungal compounds, are used to treat animals,
CC especially humans, infected with chitin containing parasites such as
CC fungi. Fungal infection treated include candidiasis, aspergillosis,
CC coccidioidomycosis, blastomycosis, paracoccidioidomycosis,
CC mucormycosis, histoplasmosis, cryptococcosis, chromoblastomycosis
CC sporotrichosis, and dermatophytoses.

```
CC   Chitin can be degraded by the enzyme chitinase. Use of whole chitinase
CC   protein for treating infections, especially fungal infections, is
CC   problematic. In view of the increasing incidents of life threatening
CC   fungal infection in e.g. immunocompromised individuals, there exists a
CC   need for identifying new compounds for treating fungal infection. The
CC   chitin-binding fragments of the present invention provide this need.
XX
SQ   Sequence   423 AA;


  Query Match             87.7%;  Score 57;  DB 20;  Length 423;
  Best Local Similarity   91.7%;  Pred. No. 0.01;
  Matches   11;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 PTFGRSFTLASS 12
           || : ||||||:
Db     244 PTYGRSITLASS 255


RESULT  15
AAW08585
ID   AAW08585 standard; Protein; 387 AA.
XX
AC   AAW08585;
XX
DT   24-MAR-1997  (first entry)
XX
DE   Human 39 kDa chitinase.
XX
KW   Chitinase; chitotriosidase; chitin; infections disease;
KW   gene therapy; vaccine; diagnosis; lysosomal lipidosis;
KW   Gaucher disease; leishmaniasis; sarcoidosis; multiple sclerosis;
KW   X linked adrenoleukodystrophy; drug delivery; cosmetics; food.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Peptide         1..21
FT                   /label= Sig_peptide
FT   Domain          178..198
FT                   /label= Catalytic_domain
XX
PN   WO9640940-A2.
XX
PD   19-DEC-1996.
XX
PF   06-JUN-1996;   96WO-NL00225.
XX
PR   07-JUN-1995;   95US-0486839.
XX
PA   (UNAM-) UNIV AMSTERDAM.
XX
IN   Aerts JMFG;
XX
DR   WPI; 1997-118698/11.
DR   N-PSDB; AAT50834.
XX
PT   New human chitinase - used to treat or prevent infection by
PT   chitin-containg pathogens, in diagnosis and as additives to cosmetics,
PT   foods, implants etc.
XX
PS   Claim 1; Page 43-45; 58pp; English.
XX
CC   Human chitinases of 50 kDa (AAW08584) and 39 kDa (AAW08585) are stable
CC   to many proteases, active at pH 3-8 and up to 50 deg. and stable in
CC   the circulation. They are the result of alternative splicing of
CC   RNA, and can be produced on a large-scale in transformed host cells
CC   using cDNA clones (see also AA15083-84) obtd. from a human macrophage
CC   library. The chitinases, optionally expressed from a gene therapy
CC   vector, are used to treat or prevent infection by chitin-containing
CC   pathogens (e.g. fungi, protozoa, helminths) and more generally to
CC   degrade chitin. They are also useful in controlled release drug
CC   delivery, in cosmetics, foods and dental products, for antibody
CC   prodn. and for diagnosis of diseases associated with elevated
```

RC chitinase levels.
XX
SQ Sequence 487 AA;

Query Match 87.7%; Score 57; DB 18; Length 487;
Best Local Similarity 91.7%; Pred. No. 0.01;
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FIFKSFILASS 12
||||||||||| |
Db 265 pfyrsfilass 276

Search completed: August 26, 2002, 06:13:25
Job time: 247 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 26, 2002, 06:23:28 ; Search time 6.22 Seconds
(without alignments)
6.870 Million cell updates/sec

Title: US-09-744-282-5
Perfect score: 65
Sequence: 1 PIFGRSFILASSE 13

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 253628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 253628

Minimum DB seq length: 0
Maximum DB seq length: 20000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep: *
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep: *
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep: *
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep: *
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep: *
6: /cgn2_6/ptodata/1/iaa/backfile1.pep: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	65	100.0	13	1	US-08-619-645-5	Sequence 5, Appli
2	65	100.0	13	2	US-08-644-493-5	Sequence 5, Appli
3	65	100.0	16	4	US-09-171-705-40	Sequence 40, Appl
4	57	87.7	373	4	US-09-039-198A-14	Sequence 14, Appl
5	57	87.7	373	4	US-09-039-198A-15	Sequence 15, Appl
6	57	87.7	387	2	US-08-486-839-6	Sequence 6, Appli
7	57	87.7	387	3	US-09-151-011-6	Sequence 6, Appli
8	57	87.7	387	4	US-09-343-623-6	Sequence 6, Appli
9	57	87.7	466	2	US-08-486-839-4	Sequence 4, Appli
10	57	87.7	466	3	US-09-151-011-4	Sequence 4, Appli
11	57	87.7	466	4	US-09-039-198A-2	Sequence 2, Appli
12	57	87.7	466	4	US-09-039-198A-4	Sequence 4, Appli
13	57	87.7	466	4	US-09-343-623-4	Sequence 4, Appli
14	54	83.1	385	2	US-08-694-915-2	Sequence 2, Appli
15	54	83.1	416	2	US-08-694-915-4	Sequence 4, Appli
16	47	72.3	16	4	US-09-171-705-41	Sequence 41, Appl
17	44	67.7	9	1	US-08-619-645-1	Sequence 1, Appli
18	44	67.7	9	2	US-08-644-493-1	Sequence 1, Appli
19	43	66.1	13	1	US-08-619-645-6	Sequence 6, Appli
20	43	66.1	13	2	US-08-644-493-6	Sequence 6, Appli
21	43	66.1	98	1	US-07-612-674-12	Sequence 12, Appl
22	43	66.1	276	1	US-07-612-674-5	Sequence 5, Appli
23	43	66.1	279	1	US-07-612-674-8	Sequence 8, Appli
24	40	61.5	554	2	US-08-524-051-2	Sequence 2, Appli
25	40	61.5	554	3	US-09-052-778-16	Sequence 16, Appl
26	39	60.0	16	4	US-09-171-705-39	Sequence 39, Appl
27	38	58.8	198	4	US-09-080-983-15	Sequence 15, Appl
28	34	52.3	1012	1	US-08-219-262B-10	Sequence 10, Appl
29	34	52.3	1012	3	US-09-031-655-10	Sequence 10, Appl
30	34	52.3	1013	2	US-08-708-543A-34	Sequence 34, Appl
31	33	50.8	147	4	US-09-166-637A-40	Sequence 40, Appl
32	33	50.8	218	3	US-09-239-011-4	Sequence 4, Appl
33	33	50.8	218	4	US-08-878-862-4	Sequence 4, Appli
34	33	50.8	384	2	US-08-637-759B-375	Sequence 375, Appl
35	33	50.8	384	3	US-08-871-455A-375	Sequence 375, Appl
36	33	50.8	384	4	US-09-201-945-375	Sequence 375, Appl
37	32	49.2	9	1	US-08-619-645-2	Sequence 2, Appli
38	32	49.2	9	2	US-08-644-493-2	Sequence 2, Appli
39	32	49.2	157	4	US-09-162-021B-10	Sequence 10, Appl
40	32	49.2	197	4	US-09-162-021B-8	Sequence 8, Appli
41	32	49.2	198	4	US-09-162-021B-6	Sequence 6, Appli
42	32	49.2	384	4	US-08-857-087-104	Sequence 104, Appl
43	32	49.2	391	5	PCT-US95-15600-2	Sequence 2, Appli
44	32	49.2	568	1	US-07-742-242-3	Sequence 3, Appli
45	32	49.2	1027	4	US-09-162-021B-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1
US-08-619-645-5
; Sequence 5, Application US/08619645
; Patent No. 5736507
; GENERAL INFORMATION:
; APPLICANT: Boots, Anna M.H.
; APPLICANT: Verheijden, Gijsbertus F.M.
; TITLE OF INVENTION: No. 57365072 peptides derived from
; TITLE OF INVENTION: Preparation for use in immunotherapy of autoimmune
; TITLE OF INVENTION: Disease
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Akzo Nobel Inc. 5736507o1 Patent Department
; STREET: 1300 Piccard Drive, Suite 206
; CITY: Rockville
; STATE: Maryland
; COUNTRY: U.S.A.
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25(EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/619,645
; FILING DATE: 25-MAR-1996
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Gormley, Mary E.
; REGISTRATION NUMBER: 34,409
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 947-4433
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 13 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-619-645-5

Query Match 100.0%; Score 65; DB 1; Length 13;
Best Local Similarity 100.0%; Pred. No. 1.4e-07;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PIFGRSFILASSE 13
||||||||||||| |
Db 1 PIFGRSFILASSE 13

```
RESULT
US-09-744-282-5
  Query Match  Application US/09644044
  Patent N  6413439
  GENERAL INFORMATION:
    APPLICANT:  A.M.H. Boots
    APPLICANT:  E.E.M. Verheijden
    APPLICANT:  B.J. Bos
    TITLE OF INVENTION:  No. 9644044 el Peptides Derived from Antibodies Directed to
    TITLE OF INVENTION:  in Immunotherapy of Allergic Diseases
    NUMBER OF SEQUENCES:  10
    CORRESPONDENT ADDRESS:
      ADDRESSEE:  Akzo No. 9644044el Patent Department
      STREET:  Po Box 20
      CITY:  P
      STATE:  Maryland
      COUNTRY:  USA
      ZIP:
    COMPUTER READABLE FORM:
      MEDIUM TYPE:  Floppy disk
      COMPUTER:  IBM PC compatible
      OPERATING SYSTEM:  PC-Dos/MS-Dos
      SOFTWARE:  Patentin Release #1.0, Version #1.25
    CURRENT APPLICATION DATA:
      APPLICATION NUMBER:  US/09/744,282
      FILING DATE:  18-APR-1996
    PRIOR APPLICATION DATA:
      APPLICATION NUMBER:  US 08/619,645
      FILING DATE:  25-MAR-1996
    PRIOR APPLICATION DATA:
      APPLICATION NUMBER:  PCT/EP95/04101
      FILING DATE:  25-OCT-1995
    PRIOR APPLICATION DATA:
      APPLICATION NUMBER:  NL 9402087
      FILING DATE:  27-OCT-1994
    PRIOR APPLICATION DATA:
      APPLICATION NUMBER:  NL 9500860
      FILING DATE:  27-APR-1995
    TELECOMMUNICATION INFORMATION:
      TELEPHONE:  (+1) 256-5200
      TELEFAX:  (+1) 977-0847
  INFORMATION FOR SEQ ID NO:  5:
    SEQUENCE CHARACTERISTICS:
      LENGTH:  19 amino acids
      TYPE:  amino acid
      STRANDEDNESS:  single
      TOPOLOGY:  linear
    MOLECULE TYPE:  peptide
US-09-744-282-5

  Query Match  100.0%;  Score 95;  DB 2;  Length 19;
  Best Local Similarity  100.0%;  Pred. No. 1.9e-06;
  Matches  19;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy       1 PIEGRSFILASSE 13
           |||||||||||||
Db       1 PIEGRSFILASSE 13


RESULT
US-09-171-205-1
; Sequence 1, Application US/09171205
; Patent No. 6242208
; GENERAL INFORMATION:
;  APPLICANT: DE VOS, ANNA M.H.
;  APPLICANT: VERHEIJDEN, GIJSBERTUS F.M.
;  TITLE OF INVENTION: NOVEL PEPTIDES SUITABLE FOR USE IN ANTIGEN SPECIFIC
;  TITLE OF INVENTION: IMMUNOSUPPRESSIVE THERAPY
;  FILE REFERENCE:  7.0198 US
;  CURRENT APPLICATION NUMBER: US/09/171,205
;  CURRENT FILING DATE: 1999-02-09
;  NUMBER OF SEQ ID NOS: 78
;  SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 40
;  LENGTH: 16
;  TYPE: PRT
;  ORGANISM: Artificial Sequence
;  FEATURE:
;  OTHER INFORMATION: Description of Artificial Sequence: DERIVED FROM
;  OTHER INFORMATION: SEQUENCE OF HUMAN CARTILAGE (BC)  /note="
US-09-171-205-40

  Query Match  100.0%;  Score 95;  DB 4;  Length 16;
  Best Local Similarity  100.0%;  Pred. No. 2.4e-06;
  Matches  19;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy       1 PIEGRSFILASSE 13
           |||||||||||||
Db       4 PIEGRSFILASSE 16


RESULT  4
US-09-039-198A-14
; Sequence 14, Application US/09039198A
; Patent No. 6200951
; GENERAL INFORMATION:
;  APPLICANT:  Gray, Patrick W.
;  APPLICANT:  Tjoelker, Larry W.
;  TITLE OF INVENTION:  CHITINASE CHITIN BINDING FRAGMENTS
;  NUMBER OF SEQUENCES:  34
;  CORRESPONDENCE ADDRESS:
;  ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray & Borun
;  STREET:  233 South Wacker Drive/6300 Sears Tower
;  CITY:  Chicago
;  STATE:  Illinois
;  COUNTRY:  United States of America
;  ZIP:  60606-6402
;  COMPUTER READABLE FORM:
;  MEDIUM TYPE:  Floppy disk
;  COMPUTER:  IBM PC compatible
;  OPERATING SYSTEM:  PC-DOS/MS-DOS
;  SOFTWARE:  Patentin Release #1.0, Version #1.30
;  CURRENT APPLICATION DATA:
;  APPLICATION NUMBER:  US/09/039,198A
;  FILING DATE:
;  CLASSIFICATION:  435
;  ATTORNEY/AGENT INFORMATION:
;  NAME:  Kin-Lauros, Li Hsien
;  REGISTRATION NUMBER:  33,547
;  REFERENCE/DOCKET NUMBER:  27866/34.0g1
;  TELECOMMUNICATION INFORMATION:
;  TELEPHONE:  (312) 474-6300
;  TELEFAX:  (412) 474-0448
;  INFORMATION FOR SEQ ID NO:  14:
;  SEQUENCE CHARACTERISTICS:
;  LENGTH:  873 amino acids
;  TYPE:  amino acid
;  STRANDEDNESS:  single
;  TOPOLOGY:  linear
;  MOLECULE TYPE:  peptide
US-09-039-198A-14

  Query Match  87.7%;  Score 57;  DB 4;  Length 873;
  Best Local Similarity  91.7%;  Pred. No. 0.0031;
  Matches  11;  Conservative  1;  Mismatches  0;  Indels  0;  Gaps  0;

Qy       1 PIEGRSFILASS 12
           ||:|||||||||
Db     244 PIYGRSFILASS 255


RESULT  5
```

```
US-09-030-198A-15
; Sequence 15, Application US/09030198A
; Patent No. 6200951
;  GENERAL INFORMATION:
;    APPLICANT:  Gay, Patrick W.
;    APPLICANT:  Doolker, Larry W.
;    TITLE OF INVENTION:  CHITINASE CHITIN BINDING FRAGMENTS
;    NUMBER OF SEQUENCES:  24
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray & Borun
;      STREET:  233 South Wacker Drive/6300 Sears Tower
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  United States of America
;      ZIP:  60606-6402
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC DOS/MS DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/030,198A
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Rinehart, Blaine
;      REGISTRATION NUMBER:  33,547
;      REFERENCE/DOCKET NUMBER:  27866/34391
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (312) 474-6300
;      TELEFAX:  (312) 474-0448
;  INFORMATION FOR SEQ ID NO:  15:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  373 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-09-030-198A-15


  Query Match             87.7%;  Score 57;  DB 4;  Length 373;
  Best Local Similarity   91.7%;  Pred. No. 0.0041;
  Matches   11;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 PTFGRSFTLASS 12
           ||:|||||||||
Db     244 PTYGRSFTLASS 255


RESULT   6
US-08-486-839-6
; Sequence 6, Application US/08486839
; Patent No. 5798928
;  GENERAL INFORMATION:
;    APPLICANT:
;    TITLE OF INVENTION:  A human chitinase, its recombinant
;    TITLE OF INVENTION:  production, its use for decomposing chitin, its use
;    TITLE OF INVENTION:  in therapy or prophylaxis against infection diseases.
;    NUMBER OF SEQUENCES:  16
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Bottmann & Baron
;      STREET:  350 Jericho Turnpike
;      CITY:  Jericho
;      STATE:  New York
;      COUNTRY:  United States of America
;      ZIP:  11758
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC DOS/MS DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;      SOFTWARE:
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/486,839
;      FILING DATE:  07 - June - 1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Baron, Ronald J.
;      REGISTRATION NUMBER:  29,281
;      REFERENCE/DOCKET NUMBER:  294 26
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (516) 822-4550
;      TELEFAX:  (516) 822 3582
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  387 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  unknown
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:  protein
;    HYPOTHETICAL:  NO
US-08-486-839-6


  Query Match             87.7%;  Score 57;  DB 2;  Length 387;
  Best Local Similarity   91.7%;  Pred. No. 0.0042;
  Matches   11;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 PTFGRSFTLASS 12
           ||:|||||||||
Db     265 PTYGRSFTLASS 276


RESULT   7
US-09-151-011-6
; Sequence 6, Application US/09151011
; Patent No. 6057142
;  GENERAL INFORMATION:
;    APPLICANT:
;    TITLE OF INVENTION:  A Human Chitinase, Its Recombinant
;    TITLE OF INVENTION:  Production, Its Use For Decomposing Chitin, Its Use In
;    TITLE OF INVENTION:  Therapy or Prophylaxis Against Infection Diseases.
;    NUMBER OF SEQUENCES:  16
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Bottmann & Baron, LLP
;      STREET:  6900 Jericho Turnpike
;      CITY:  Syosset
;      STATE:  New York
;      COUNTRY:  United States of America
;      ZIP:  11791
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC DOS/MS DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;      SOFTWARE:
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/151,011
;      FILING DATE:  10 - September - 1998
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Morriss, Robert C.
;      REGISTRATION NUMBER:  42,910
;      REFERENCE/DOCKET NUMBER:  294 72 US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (516) 822 4550
;      TELEFAX:  (516) 822 3582
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  387 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  unknown
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:  protein
;    HYPOTHETICAL:  NO
US-09-151-011-6
```

Query Match 87.7%; Score 57; DB 4; Length 466;
Best Local Similarity 91.7%; Pred. No. 0.0042;
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PTFGRSPTLASS 12
|:|||||||||||
Db 265 PTYGRSPTLASS 276

RESULT 10
US-09-151-011-4
; Sequence 4, Application US/09151011
; Patent No. 6057142
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: A Human Chitinase, Its Recombinant
; TITLE OF INVENTION: Production, Its Use For Decomposing Chitin, Its Use In
; TITLE OF INVENTION: Therapy Or Prophylaxis Against Infection Diseases.
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hoffmann & Baron, LLP
; STREET: 6900 Jericho Turnpike
; CITY: Syosset
; STATE: New York
; COUNTRY: United States of America
; ZIP: 11791
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC DOS/MS DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; SOFTWARE:
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/151,011
; FILING DATE: 10-September-1998
; ATTORNEY/AGENT INFORMATION:

```
;       NAME:  Merriss, Robert C.
;       REGISTRATION NUMBER:  42,910
;       REFERENCE/DOCKET NUMBER:  294-32 DIV
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (516) 822-4550
;       TELEFAX:  (516) 822-4568
;   INFORMATION FOR SEQ ID NO:  4:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  456 amino acids
;       TYPE:  amino acid
;       STRANDEDNESS:  both
;       TOPOLOGY:  unknown
;     MOLECULE TYPE:  protein
;     HYPOTHETICAL:  N
US-09-151-911-4


  Query Match             87.7%;  Score 57;  DB 4;  Length 456;
  Best Local Similarity   91.7%;  Pred. No. 0.004;
  Matches   11;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 PIFGRSFILASS 12
           ||:|||||||||
Db     265 PIYGRSFILASS 276


RESULT  11
US-09-039-198A-2
; Sequence 2, Application US/09039198A
; Patent No. 6200951
; GENERAL INFORMATION:
;     APPLICANT:  Gray, Patrick W.
;     APPLICANT:  Tjoelker, Larry W.
;     TITLE OF INVENTION:  CHITINASE CHITIN BINDING FRAGMENTS
;     NUMBER OF SEQUENCES:  34
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray & Borun
;       STREET:  233 South Wacker Drive,6300 Sears Tower
;       CITY:  Chicago
;       STATE:  Illinois
;       COUNTRY:  United States of America
;       ZIP:  60606-6402
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.30
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/09/039,198A
;       FILING DATE:
;       CLASSIFICATION:  435
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Rinehart, Kay
;       REGISTRATION NUMBER:  35,047
;       REFERENCE/DOCKET NUMBER:  27866/44791
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (312) 474-6300
;       TELEFAX:  (312) 474-0448
;   INFORMATION FOR SEQ ID NO:  2:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  456 amino acids
;       TYPE:  amino acid
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  protein
US-09-039-198A-2


  Query Match             87.7%;  Score 57;  DB 4;  Length 469;
  Best Local Similarity   91.7%;  Pred. No. 0.004;
  Matches   11;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 PIFGRSFILASS 12
           ||:|||||||||
Db     265 PIYGRSFILASS 276


RESULT  12
US-09-039-198A-4
; Sequence 4, Application US/09039198A
; Patent No. 6200951
; GENERAL INFORMATION:
;     APPLICANT:  Gray, Patrick W.
;     APPLICANT:  Tjoelker, Larry W.
;     TITLE OF INVENTION:  CHITINASE CHITIN BINDING FRAGMENTS
;     NUMBER OF SEQUENCES:  34
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray & Borun
;       STREET:  233 South Wacker Drive,6300 Sears Tower
;       CITY:  Chicago
;       STATE:  Illinois
;       COUNTRY:  United States of America
;       ZIP:  60606-6402
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.30
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/09/039,198A
;       FILING DATE:
;       CLASSIFICATION:  435
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Rinehart, Li-Hsien
;       REGISTRATION NUMBER:  35,047
;       REFERENCE/DOCKET NUMBER:  27866/44791
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (312) 474-6300
;       TELEFAX:  (312) 474-0448
;   INFORMATION FOR SEQ ID NO:  4:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  456 amino acids
;       TYPE:  amino acid
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  protein
US-09-039-198A-4


  Query Match             87.7%;  Score 57;  DB 4;  Length 469;
  Best Local Similarity   91.7%;  Pred. No. 0.004;
  Matches   11;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 PIFGRSFILASS 12
           ||:|||||||||
Db     265 PIYGRSFILASS 276


RESULT  13
US-09-443-623-4
; Sequence 4, Application US/09443623
; Patent No. 6303118
; GENERAL INFORMATION:
;     APPLICANT:
;     TITLE OF INVENTION:  A human chitinase, its recombinant
;     TITLE OF INVENTION:  production, its use to decompose chitin, its use
;     TITLE OF INVENTION:  in therapy or prophylaxis against infectious diseases,
;     NUMBER OF SEQUENCES:  16
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Hoffmann & Baron
;       STREET:  350 Jericho Turnpike
;       CITY:  Jericho
;       STATE:  New York
;       COUNTRY:  United States of America
;       ZIP:  11758
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
```

```
;       OPERATING SYSTEM:  IBM DOS/MS DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.30
;       SOFTWARE:
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/241,625
;       FILING DATE:
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/199,829
;       FILING DATE:  07-June-1995
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Barr, Ronald L.
;       REGISTRATION NUMBER:  28,291
;       REFERENCE/DOCKET NUMBER:  204.26
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (310) 822-9660
;       TELEFAX:  (310) 822-9679
;   INFORMATION FOR SEQ ID NO:  4:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  [illegible] amino acids
;       TYPE:  amino acid
;       STRANDEDNESS:  unknown
;       TOPOLOGY:  unknown
;     MOLECULE TYPE:  protein
;     HYPOTHETICAL:  NO
US-08-694-915-4

  Query Match             87.7%;  Score 57;  DB 4;  Length 485;
  Best Local Similarity   91.7%;  Pred. No. ....;
  Matches   22;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      [illegible]
        [illegible]
Db    262 [illegible]


RESULT  [illegible]
US-08-694-915-4
; Sequence 4, Application US/08694915
; Patent No. [illegible]
;  GENERAL INFORMATION:
;    APPLICANT:  [illegible]
;    APPLICANT:  [illegible]
;    APPLICANT:  [illegible]
;    TITLE OF INVENTION:  [illegible]
;    NUMBER OF SEQUENCES:  [illegible]
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SmithKline Beecham Corporation
;      STREET:  709 Swedeland Road
;      CITY:  King of Prussia
;      STATE:  PA
;      COUNTRY:  USA
;      ZIP:  19406-2799
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  [illegible]
;      COMPUTER:  [illegible]
;      OPERATING SYSTEM:  [illegible]
;      SOFTWARE:  [illegible]
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/694,915
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  [illegible]
;      REGISTRATION NUMBER:  [illegible]
;      REFERENCE/DOCKET NUMBER:  [illegible]
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  [illegible]
;      TELEFAX:  [illegible]
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  [illegible]:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  [illegible] amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  [illegible]
;      TOPOLOGY:  [illegible]
;    MOLECULE TYPE:  [illegible]
;    HYPOTHETICAL:  [illegible]
;    ANTI-SENSE:  [illegible]
;    FRAGMENT TYPE:  [illegible]
;    ORIGINAL SOURCE:
US-08-694-915-[illegible]
```

Query Match 83.1%; Score 54; DB 2; Length 116;
Best Local Similarity 76.9%; Prod. No. 0.01;
Matches 10; Conservative 2; Mismatches 1; Indels 2; Gaps 1;

Qy 1 FIFHSFIIASSE 14
| | | | | |
Db 293 FIYHSFIIASAE 305

Search completed: August 26, 2002, 06:23:29
Job time: 87 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

SW protein protein search, using sw model

Run on: August 26, 2002, 06:10:39 ; Search time 7.5 Seconds
(without alignments)
11.702 Million cell updates/sec

Title: US-09-744-282-1
Perfect score: 44
Sequence: 1 FGRSFFLAS 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 28518 seqs, 9608944 residues

Total number of hits satisfying chosen parameters: 28518

Minimum DB seq length: 0
Maximum DB seq length: 9

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 71:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	18	40.9	8	2	A37521	R-phycoerythrin ga
2	17	38.6	7	2	A25269	sex pheromone cAM3
3	17	38.6	9	2	S39766	cardioactive pepti
4	16	36.4	9	2	PI0324	Ig heavy chain CRD
5	15	34.1	5	2	PT0644	T cell receptor be
6	15	34.1	8	2	A23967	leucopyrokinin - M
7	15	34.1	9	2	A26563	ard, active pept
8	15	34.1	9	2	S27233	ard, active pepti
9	15	34.1	9	2	S39767	Card, active pepti
10	14	29.5	8	2	T10952	hypothetical prote
11	14	29.5	8	2	A39892	F-element, P-cytot
12	14	29.5	9	2	S02384	probable membrane
13	14	29.5	9	2	E41978	calliFMRFamide 5
14	14	29.5	9	2	F41978	calliFMRFamide 6
15	14	29.5	9	2	G41978	calliFMRFamide 7 -
16	12	27.3	5	2	A44955	alkanal monooxygen
17	12	27.3	5	2	PT0278	Ig heavy chain CRD
18	12	27.3	5	2	PI0553	T cell receptor be
19	12	27.3	5	2	S69237	surface protein %e
20	12	27.3	7	1	A61324	dermorphin - Rohde
21	12	27.3	7	2	S36662	dermorphin (Lys 7)
22	12	27.3	7	2	I40504	hypothetical prote
23	12	27.3	8	2	S08996	hypothetical semic
24	12	27.3	8	2	S10596	adipokinetic hormo
25	12	27.3	8	2	B49823	adipokinetic hormo
26	12	27.3	8	2	B44960	beta-peptide isofo
27	12	27.3	8	2	A33995	adipokinetic hormo
28	12	27.3	8	2	S55810	adipokinetic hormo
29	12	27.3	8	2	A58620	adipokinetic hormo
30	12	27.3	8	2	I48699	Ig: ethereal proto
31	12	27.3	8	2	D40701	unidentified 67.7k
32	12	27.3	8	2	S11078	d-, ose 6-phosphat
33	12	27.3	8	2	PI0162	paramyosin - horse
34	12	27.3	8	2	JS0215	le-sokinin V - Mad
35	12	27.3	8	2	I47493	neuropeptide calli
36	12	27.3	8	2	E47493	neuropeptide calli
37	12	27.3	8	2	PH0855	T cell receptor al
38	12	27.3	8	2	B74574	Fe-meet-Y and beta
39	12	27.3	9	2	A24244	ast pokinetic horm
40	12	27.3	9	2	F28854	tibiotropic peptid
41	12	27.3	9	2	I24180	fibrinogen beta ch
42	12	27.3	9	2	S15855	vitamin D-25-hydro
43	12	27.3	9	2	B45626	probable nitrogeny
44	12	27.3	9	2	G58502	K, heavy and bivalent
45	12	27.3	9	2	B25569	serum amyloid P-c

ALIGNMENTS

RESULT 1
A37521
R-phycoerythrin gamma-E chain - red alga (Gastroclonium coulteri) (fragment)
C;Species: Gastroclonium coulteri
C;Date: 07-Mar-1988 #sequence_revision 07-Mar-1988 #text_change 16-Jul-1999
C;Accession: A37521; J22565
R;Klotz, A.V.; Glazer, A.N.
J. Biol. Chem. 260, 4856-4863, 1985
A;Title: Characterization of the linin attachment sites in R-phycoerythrin.
A;Reference number: A22565; MUID:85182601
A;Accession: A37521
A;Molecule type: protein
A;Residues: 1-8 <KLO>

Query Match 40.9%; Score 18; DB 2; Length 8
Best Local Similarity 80.0%; Pred. No. 2.3e-05;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 4 SFFLA 8
|| ||
Db 2 SFALA 6

RESULT 2
A25269
sex pheromone cAM373 - Enterococcus faecalis
N;Alternate names: clumping-inducing agent ('IA)
C;Species: Enterococcus faecalis
C;Date: 31-Mar-1988 #sequence_revision 31-Mar-1988 #text_change 16-Jul-1999
C;Accession: A25269
R;Mori, M., Tanaka, H., Sakagami, Y., Isogai, A., Fujita, M. & Kitada, C.; Whoto, S.A.
FEBS Lett. 206, 69-72, 1986
A;Title: Isolation and structure of the Streptococcus faecalis sex pheromone, cAM373.
A;Reference number: A25269; MUID:87005252
A;Accession: A25269
A;Molecule type: protein
A;Residues: 1-7 <MOR>

Query Match 38.6%; Score 17; DB 2; Length 7;
Best Local Similarity 80.0%; Pred. No. 3.9e-05;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 5 FFLAS 9
| |||
Db 4 FILAS 7

RESULT 3
S39766
cardioactive peptide CCAP - yellow mealworm

C;Species: Tenebrio molitor (yellow mealworm)
C;Date: [illegible] #sequence_revision 12-Apr-1996 #text_change 07-May-1996
C;Accession: [illegible]
R;Kempe, K.; Stern, S.; Reynolds, S.E.; Stay, B.; Hackett, M.; Schooley, D.A.
Biochem. Biophys. Res. Commun. 174, 1069-1074, 1991
A;Title: Isolation and identification of a cardioactive peptide from Tenebrio molitor an
A;Reference number: [illegible]; MUID:91172032
A;Accession: [illegible]
A;Molecule type: protein
A;Residues: [illegible]
C;Keywords: [illegible] neuropeptide

Query Match 88.6%; Score 17; DB 2; Length 9;
Best Local Similarity 50.0%; Prod. No. 2.8e-05;
Matches 4; Conservative 1; Mismatches 2; Indels 1; Gaps 0;

Qy 1 FDRSFI 6
| |
Db 1 FXNAFI 7

RESULT 4
PI0324
leucokinin I - cockroach (Periplaneta americana)
C;Species: Periplaneta americana (American cockroach)
C;Date: 01-Jan-1988 #sequence_revision 01-Jan-1988 #text_change 09-Jul-1999
C;Accession: PI0324
R;Holman, G.M.; Cook, B.J.; Nachman, R.J.
Comp. Biochem. Physiol., C 84, 423-428, 1986
A;Title: Primary structure of leucokinin I from the American cockroach, Periplaneta ame
A;Reference number: PI0222; MUID:87118337
A;Accession: PI0324
A;Molecule type: protein
A;Residues: 1-9 <HOL>
C;Keywords: amidated carboxyl end; neuropeptide; pyroglutamic acid
F;1/Modified site: pyrrolidone carboxylic acid (Glp) #status experimental
F;9/Modified site: amidated carboxyl end (Leu) #status experimental

Query Match 88.6%; Score 17; DB 2; Length 9;
Best Local Similarity 50.0%; Prod. No. 2.8e-05;
Matches 4; Conservative 1; Mismatches 2; Indels 1; Gaps 0;

Qy 2 FDRSFI 7
| |
Db 1 FXNAFI 6

RESULT 5
PI0324
leucokinin I - cockroach (Periplaneta americana)
C;Species: Periplaneta americana (American cockroach)
C;Date: 01-Jan-1988 #sequence_revision 01-Jan-1988 #text_change 09-Jul-1999
C;Accession: PI0324
R;Holman, G.M.; Cook, B.J.; Nachman, R.J.
Comp. Biochem. Physiol., C 84, 423-428, 1986
A;Title: Primary structure of leucokinin I from the American cockroach, Periplaneta ame
A;Reference number: PI0222; MUID:87118337
A;Accession: PI0324
A;Molecule type: protein
A;Residues: 1-9 <HOL>
C;Keywords: amidated carboxyl end; neuropeptide; pyroglutamic acid
F;1/Modified site: pyrrolidone carboxylic acid (Glp) #status experimental
F;9/Modified site: amidated carboxyl end (Leu) #status experimental

Query Match 88.6%; Score 17; DB 2; Length 9;
Best Local Similarity 50.0%; Prod. No. 2.8e-05;
Matches 4; Conservative 1; Mismatches 2; Indels 1; Gaps 0;

Qy 3 FDRSFI 8
| |
Db 2 FXNAFI 7

RESULT 6
A23967
leucopyrokinin - Madeira cockroach
C;Species: Leucophaea maderae (Madeira cockroach)
C;Date: 31-Mar-1988 #sequence_revision 26-May-1994 #text_change 11-Jul-1997
C;Accession: A23967
R;Nachman, R.J.; Holman, G.M.; Cook, B.J.
Biochem. Biophys. Res. Commun. 137, 938-942, 1986
A;Title: Active fragments and analogs of the insect neuropeptide leucopyrokinin: stru
A;Reference number: A23967; MUID:86269041
A;Accession: A23967
A;Molecule type: protein
A;Residues: 1-8 <NAC>
C;Keywords: amidated carboxyl end; neuropeptide; pyroglutamic acid
F;1/Modified site: pyrrolidone carboxylic acid (Glp) #status experimental
F;8/Modified site: amidated carboxyl end (Leu) #status experimental

Query Match 84.1%; Score 15; DB 2; Length 8;
Best Local Similarity 100.0%; Prod. No. 2.8e-05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 SFI 6
|||
Db 3 SFI 5

RESULT 7
A26363
cardioactive peptide - green crab (fragment)
C;Species: Carcinus maenas (green crab, common shore crab)
C;Date: 25-Oct-1987 #sequence_revision 25-Oct-1987 #text_change 18-Jun-1999
C;Accession: A26363
R;Stangier, J.; Hilbich, C.; Beyreuther, K.; Keller, R.
Proc. Natl. Acad. Sci. U.S.A. 84, 575-579, 1987
A;Title: Unusual cardioactive peptide (CCAP) from pericardial organs of the shore cra
A;Reference number: A26363
A;Accession: A26363
A;Molecule type: protein
A;Residues: 1-9 <STA>

Query Match 84.1%; Score 15; DB 2; Length 9;
Best Local Similarity 50.0%; Prod. No. 2.8e-05;
Matches 3; Conservative 1; Mismatches 2; Indels 2; Gaps 0;

Qy 1 FCRSFI 6
| :||
Db 2 FCNAFI 7

RESULT 8
S27233
cardioactive peptide CCAP - tobacco hornworm
C;Species: Manduca sexta (tobacco hornworm)
C;Date: 09-Jun-1994 #sequence_revision 12-Apr-1996 #text_change 12-Apr-1996
C;Accession: S27233
R;Cheung, C.C.; Loi, P.K.; Sylwester, A.W.; Lee, T.D.; Dublitz, N.J.
FEBS Lett. 313, 165-168, 1992
A;Title: Primary structure of a cardioactive neuropeptide from the tobacco hawkmoth,
A;Reference number: S27233; MUID:93050243
A;Accession: S27233
A;Molecule type: protein
A;Residues: 1-9 <CHE>
C;Keywords: neuropeptide

Query Match 84.1%; Score 15; DB 2; Length 9;
Best Local Similarity 50.0%; Prod. No. 2.8e-05;
Matches 3; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

```
Qy       1 FGRSFI 6
           | :|||
Db       2 FCNAFI 7


RESULT   9
S07677
cardioactive peptide CAP - Spodoptera eridania
C;Species: Spodoptera eridania
C;Date: [illegible] #sequence_revision 12-Apr-1996 #text_change 07-May-1999
C;Accession: S07677
R;Furuya, K.; Liao, S.; Reynolds, S.E.; Orr, G.R.; Hanke, M.; Schooley, D.A.
Biol. Chem. Hoppe Seyler 374, 1065-1074, 1993
A;Title: Isolation and identification of a cardioactive peptide in the Tenebrio molitor an
A;Reference number: S07666; MUID:94176032
A;Accession: S07677
A;Molecule type: protein
A;Residues: 1-9 <FUK>
C;Keywords: neuropeptide


  Query Match             34.1%;  Score 15;  DB 2;  Length 9;
  Best Local Similarity   50.0%;  Pred. No. 2.8e+04;
  Matches    3;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 FGRSFI 6
           | :|||
Db       2 FCNAFI 7


RESULT   10
I10952
hypothetical protein 1 - spring vetch
C;Species: Vicia sativa (spring vetch, tare)
C;Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 16-Jul-1999
C;Accession: I10952
R;Christiansen, A.; Hansen, A.C.; Vijn, I.; Pallisgard, N.; Larsen, K.; Yang, W.C.; Biss
submitted to the EMBL Data Library, December 1995
A;Description: A novel type of DNA binding protein interacts with a conserved sequence i
A;Reference number: Z17228
A;Accession: I10952
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-9 <CHR>
A;Cross-references: EMBL:X05705; NID:g1360633; PIR:S22823


  Query Match             29.5%;  Score 13;  DB 2;  Length 9;
  Best Local Similarity   75.0%;  Pred. No. 2.8e+05;
  Matches    3;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       6 ILAS 9
           | : |
Db       4 ILVS 7


RESULT   11
A30852
Polypeptide, P cytotype determining - fruit fly (Drosophila melanogaster) (fragment)
C;Species: Drosophila melanogaster
C;Date: 24-Jan-1992 #sequence_revision 24-Jan-1992 #text_change 16-Jul-1999
C;Accession: A30852
R;Nitasaka, E.; Mukai, T.; Yamazaki, T.
Proc. Natl. Acad. Sci. U.S.A. 84, 7605-7608, 1987
A;Title: Repressor of P elements in Drosophila melanogaster: cytotype determination by a
A;Reference number: A30852
A;Accession: A30852
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-8 <NIT>
C;Genetics:
A;Gene: FlyBase:P element
A;Cross-references: FlyBase:FBgn0004055


  Query Match             29.5%;  Score 13;  DB 2;  Length 8;
  Best Local Similarity   75.0%;  Pred. No. 2.8e+05;
  Matches    3;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 FGRS 4
           | :||
Db       5 FKRS 8


RESULT   12
S02384
probable membrane antigen (BLA - human herpesvirus 4 (transient)
C;Species: human herpesvirus 4, Epstein-Barr virus
C;Date: 01-Dec-1989 #sequence_revision 01-Dec-1989 #text_change 16-Jul-1999
C;Accession: S02384
R;Walls, D.; Gannon, F.
EMBO J. 7, 1191-1196, 1988
A;Title: The expression of novel antigens from the Epstein-Barr virus latent genome
A;Reference number: S02381; MUID:88279424
A;Accession: S02384
A;Molecule type: DNA
A;Residues: 1-9 <WAL>
A;Cross-references: EMBL:X07817; NID:g59173; PIR:A:CAA30976.1; PIR:A00915
C;Keywords: membrane protein; surface antigen


  Query Match             29.5%;  Score 13;  DB 2;  Length 9;
  Best Local Similarity   60.0%;  Pred. No. 2.8e+05;
  Matches    3;  Conservative    1;  Mismatches    1;  Indels    1;  Gaps    1;

Qy       4 SFILA 8
           | :||
Db       2 SLSLA 6


RESULT   13
E41978
calliFMRFamide 5 - bluebottle fly (Calliphora vomitoria)
C;Species: Calliphora vomitoria
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 17-Mar-1999
C;Accession: E41978
R;Dowe, R.; Johnsen, A.H.; Sowell, J.C.; Scott, A.G.; Orchard, I.; Rehfeld, C.F.; The
Proc. Natl. Acad. Sci. U.S.A. 89, 2429-2433, 1992
A;Title: Isolation, structure, and activity of Phe Met Arg Phe NH2 neuropeptides (d
A;Reference number: A41978; MUID:92196211
A;Accession: E41978
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-9 <DUV>
C;Keywords: amidated carboxyl end; neuropeptide
F;9/Modified site: amidated carboxyl end (Phe) #status experimental


  Query Match             29.5%;  Score 13;  DB 2;  Length 9;
  Best Local Similarity   50.0%;  Pred. No. 2.8e+05;
  Matches    2;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       2 GRSF 5
           |: |
Db       3 GQDF 6


RESULT   14
F41978
calliFMRFamide 6 - bluebottle fly (Calliphora vomitoria)
C;Species: Calliphora vomitoria
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 17-Mar-1999
C;Accession: F41978
R;Dowe, R.; Johnsen, A.H.; Sowell, J.C.; Scott, A.G.; Orchard, I.; Rehfeld, C.F.; The
Proc. Natl. Acad. Sci. U.S.A. 89, 2429-2433, 1992
A;Title: Isolation, structure, and activity of Phe Met Arg Phe NH2 neuropeptides (d
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 26, 2002, 06:24:54 ; Search time 44.94 Seconds
(without alignments)
7,754 Million cell updates/sec

Title: US-09-744-282-1
Perfect score: 44
Sequence: 1 FGRSFILAS 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 210

Minimum DB seq length: 0
Maximum DB seq length: 9

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	18	40.9	9	1	FARP_CALSI	P38495 callinectes
2	17	38.6	7	1	CIA_ENTFA	P11932 enterococcus
3	16	34.1	8	1	LPK_LEUMA	P13910 leucophaea
4	16	34.1	9	1	OCAP_CARMA	P08556 carcinus ma
5	16	34.1	9	1	31_NEPNE	P21816 nephrops no
6	13	29.5	9	1	FAR5_CALVO	P41860 calliphora
7	13	29.5	9	1	FAR6_CALVO	P41861 calliphora
8	13	29.5	9	1	FAR7_CALVO	P41862 calliphora
9	12	27.3	4	1	OCP1_OCTMI	P80648 octopus min
10	12	27.3	5	2	ALL4_CARMA	P08817 carcinus ma
11	12	27.3	7	1	ALL2_CARMA	P08809 carcinus ma
12	12	27.3	7	1	ALL3_CARMA	P08810 carcinus ma
13	12	27.3	7	1	ALL4_CARMA	P08811 carcinus ma
14	12	27.3	7	1	ALL5_CARMA	P08812 carcinus ma
15	12	27.3	7	1	ALL7_CYDPO	P82158 cydia pomon
16	12	27.3	7	1	FAR1_PROCL	P08809 procambarus
17	12	27.3	7	1	FAR2_PROCL	P08808 procambarus
18	12	27.3	8	1	AKH_LIBAU	P25318 libellula a
19	12	27.3	8	1	AKH_TABAT	P14595 tabanus atr
20	12	27.3	8	1	ALL2_CARMA	P08815 carcinus ma
21	12	27.3	8	1	ALL5_CARMA	P08818 carcinus ma
22	12	27.3	8	1	ALL7_CARMA	P08820 carcinus ma
23	12	27.3	8	1	ALL8_CARMA	P08821 carcinus ma
24	12	27.3	8	1	ALL1_CYDPO	P82152 cydia pomon
25	12	27.3	8	1	ALL3_CYDPO	P82154 cydia pomon
26	12	27.3	8	1	ALL4_CALVO	P41840 calliphora
27	12	27.3	8	1	ALL4_CYDPO	P82155 cydia pomon
28	12	27.3	8	1	ALL5_CALVO	P41841 calliphora
29	12	27.3	8	1	ALL5_CYDPO	P82156 cydia pomon
30	12	27.3	8	1	ALL6_CYDPO	P82157 cydia pomon
31	12	27.3	8	1	ALL7_CARMA	P08822 carcinus ma
32	12	27.3	8	1	ALL8_CARMA	P08823 carcinus ma
33	12	27.3	8	1	ALL9_CARMA	P08824 carcinus ma
34	12	27.3	8	1	CAB1_ENTFA	P11268 enterococcus
35	12	27.3	8	1	FAR5_BOMAM	P41866 bombyx mor
36	12	27.3	8	1	FAR4_BOMAM	P41867 bombyx mor
37	12	27.3	8	1	HTF2_PERAM	P01549 periplaneta
38	12	27.3	8	1	LCK5_LEUMA	P13987 leucophaea
39	12	27.3	8	1	ORMY_CRUCI	P9.455 crustacea
40	12	27.3	9	1	ALL0_CARMA	P81813 carcinus ma
41	12	27.3	9	1	ALL1_CARMA	P81814 carcinus ma
42	12	27.3	9	1	FIBB_ERYFA	P.8836 erythrocebu
43	12	27.3	9	1	FIBB_THEGE	P.8842 theropithec
44	12	27.3	9	1	SAMP_MUSCA	P.4095 castelus ca
45	11	25.0	5	1	KES1_CITKO	P20.2 kluyveromyc

ALIGNMENTS

RESULT 1
FARP_CALSI
ID FARP_CALSI STANDARD; PRT; 9 AA.
AC P38495;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE FMRFamide-like neuropeptide.
OS Callinectes sapidus (Blue crab).
OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Brachyura;
OC Eubrachyura; Portunoidea; Portunidae; Callinectes.
OX NCBI_TaxID=6763;
RN [1]
RP SEQUENCE.
RX MEDLINE=92270479; PubMed=1815216;
RA Krajniak K.G.;
RT "The identification and structure-activity relations of a
RT cardioactive FMRFamide-related peptide from the blue crab Callinectes
RT sapidus.";
RL Peptides 12:1295-1302(1991).
CC -!- FUNCTION: CARDIOACTIVE PEPTIDE.
CC -!- SIMILARITY: BELONGS TO THE FARP (FMRFAMIDE RELATED PEPTIDE)
CC FAMILY.
KW Neuropeptide; Amidation.
FT MOD_RES 9 9 AMIDATION.
SQ SEQUENCE 9 AA; 1159 MW; 134F072905A4045F CRC64;

Query Match 40.9%; Score 18; DB 1; Length 9;
Best Local Similarity 60.0%; Pred. No. 1e+05;
Matches 3; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 FGRSF 5
: |||
Db 2 YNRSF 6

RESULT 2
CIA_ENTFA
ID CIA_ENTFA STANDARD; PRT; 7 AA.
AC P11932;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 01-FEB-1991 (Rel. 17, Last annotation update)
DE Sex pheromone CAM373 (clumping inducing agent) (CIA).
OS Enterococcus faecalis (Streptococcus faecalis).
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Enterococcaceae;
OC Enterococcus.
OX NCBI_TaxID=1351;
RN [1]
RP SEQUENCE.
RX MEDLINE=87095252; PubMed=3093276;
RA Mori M., Tanaka H., Sakamura Y., Ishii A., Fujii M., Kikuchi Y.,
RA White B.A., An F.Y., Clewell D.B., Suzuki A.;
RT "Isolation and structure of the Streptococcus faecalis sex pheromone

```
RESULT   3
CCAP_CARMA
ID   CCAP_CARMA      STANDARD;      PRT;      9 AA.
AC   P38596;
DT   01-OCT-1994 (Rel. 30, Created)
DT   01-OCT-1994 (Rel. 30, Last sequence update)
DT   15-DEC-1998 (Rel. 37, Last annotation update)
DE   Cardioactive peptide (CCAP).
OS   Carcinus maenas (Common shore crab) (Green crab),
OS   Manduca sexta (Tobacco hawkmoth) (Tobacco hornworm),
OS   Tenebrio molitor (Yellow mealworm), and
OS   Spodoptera eridania (Southern armyworm).
OC   Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC   Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Brachyura;
OC   Eubrachyura; Portunoidea; Portunidae; Carcinus.
OX   NCBI_TaxID=6759, 7130, 7067, 7547;
RN   [1]
RP   SEQUENCE.
RC   SPECIES=C.maenas; TISSUE=Pericardial organs;
RA   Stanisz J., Hillisch C., Beyreuther K., Keller R.;
RT   "Neuronal cardioactive peptide (CCAP) from pericardial organs of the
RT   shore crab Carcinus maenas.";
RL   Proc. Natl. Acad. Sci. U.S.A. 84:575-579(1987).
RN   [2]
RP   SEQUENCE.
RC   SPECIES=M.sexta;
RX   MEDLINE=93050243; PubMed=1426284;
RA   Choung C.C., Loi P.K., Sylwester A.W., Lee T.D., Tublitz N.J.;
RT   "Primary structure of a cardioactive neuropeptide from the tobacco
RT   hawkmoth, Manduca sexta.";
RL   FEBS Lett. 313:165-168(1992).
RN   [3]
RP   SEQUENCE.
RC   SPECIES=T.molitor, and S.eridania; TISSUE=Head;
RX   MEDLINE=94176032; PubMed=8129851;
RA   Furuya K., Liao S., Reynolds S.E., Ota R.B., Beckett M.,
RA   Schooley D.A.;
RT   "Isolation and identification of a cardioactive peptide from Tenebrio
RT   molitor and Spodoptera eridania.";
RL   Biol. Chem. Hoppe-Seyler 374:1065-1074(1993).
CC   -!- FUNCTION: THE EFFECT OF CCAP IS BOTH IN- AND CHRONOTROPIC.
CC   -!- TISSUE SPECIFICITY: STORED IN PERICARDIAL ORGANS AND RELEASED
CC       INTO THE HEMOLYMPH.
DR   PIR; A26363; A26363.
DR   PIR; S27233; S27233.
KW   Neuropeptide; Amidation.
FT   DISULFID      3      9
FT   MOD_RES       9      9       AMIDATION.
SQ   SEQUENCE   9 AA;  959 MW;  C5A861A9C0D44EBC CRC64;


  Query Match             51.1%;  Score 15;  DB 1;  Length 9;
  Best Local Similarity   50.0%;  Prod. No. 1e+05;
  Matches    3;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 FGRSFI 6
           | :||
Db       2 FGNAFI 7


RESULT   4
DI_NEPNO
ID   DI_NEPNO        STANDARD;      PRT;      9 AA.
AC   P24866;
DT   02-MAR-1992 (Rel. 21, Created)
DT   01-MAR-1992 (Rel. 21, Last sequence update)
DT   01-MAR-1992 (Rel. 21, Last annotation update)
DE   Gastrin/cholecystokinin like peptide DI.
OS   Nephrops norvegicus (Norway lobster).
OC   Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC   Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Astacidea;
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 26, 2002, 06:10:39 ; Search time 12.4 Seconds
(without alignments)
7.903 Million cell updates/sec

Title: US-09-744-282-1
Perfect score: 44
Sequence: 1 FGRSFHLAS 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 742574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 742574

Minimum DB seq length: 0
Maximum DB seq length: 9

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_032802:*
1: /SIDS5/gcgdata/geneseq/geneseq-aa1:AA*:*
2: /SIDS5/gcgdata/geneseq/geneseq-aa1:AB*:*
3: /SIDS5/gcgdata/geneseq/geneseq-aa1:AC*:*
4: /SIDS5/gcgdata/geneseq/geneseq-aa1:AD*:*
5: /SIDS5/gcgdata/geneseq/geneseq-aa1:AE*:*
6: /SIDS5/gcgdata/geneseq/geneseq-aa1:AF*:*
7: /SIDS5/gcgdata/geneseq/geneseq-aa1:AG*:*
8: /SIDS5/gcgdata/geneseq/geneseq-aa1:AH*:*
9: /SIDS5/gcgdata/geneseq/geneseq-aa1:AI*:*
10: /SIDS5/gcgdata/geneseq/geneseq-aa1:AJ*:*
11: /SIDS5/gcgdata/geneseq/geneseq-aa1:AK*:*
12: /SIDS5/gcgdata/geneseq/geneseq-aa1:AL*:*
13: /SIDS5/gcgdata/geneseq/geneseq-aa1:AM*:*
14: /SIDS5/gcgdata/geneseq/geneseq-aa1:AN*:*
15: /SIDS5/gcgdata/geneseq/geneseq-aa1:AO*:*
16: /SIDS5/gcgdata/geneseq/geneseq-aa1:AP*:*
17: /SIDS5/gcgdata/geneseq/geneseq-aa1:AQ*:*
18: /SIDS5/gcgdata/geneseq/geneseq-aa1:AR*:*
19: /SIDS5/gcgdata/geneseq/geneseq-aa1:AS*:*
20: /SIDS5/gcgdata/geneseq/geneseq-aa1:AT*:*
21: /SIDS5/gcgdata/geneseq/geneseq-aa1:AU*:*
22: /SIDS5/gcgdata/geneseq/geneseq-aa1:AV*:*

Prob. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	44	100.0	9	17	AAR95071	Human cartilage gl
2	44	100.0	9	21	AAY69226	Peptide fragment o
3	28	52.3	9	19	AAW57025	Enzyme inhibitor p
4	28	52.3	9	17	AAR95072	Human cartilage gl
5	28	52.3	9	21	AAY69227	Peptide fragment o
6	28	52.3	9	21	AAY77399	Human chondrocyte
7	22	50.0	7	20	AAW92347	Latent fibroblast
8	22	50.0	7	22	AAG63003	Complementarity de
9	22	50.0	9	20	AAY10406	T cell epitope/MHC
10	22	50.0	9	21	AAB34002	Human secreted pro
11	21	47.7	4	17	AAW77483	Tetrapeptide useful
12	21	47.7	5	20	AAY83142	Human umbilical co
13	21	47.7	5	21	AAY67476	S. clava antimicro
14	21	47.7	8	17	AAY16509	Peptide sequence of
15	21	47.7	8	19	AAW40528	Marker protein of
16	21	47.7	9	16	AAB86368	HIV-1 protease inh
17	21	47.7	9	17	AAW49256	Human leucocyte an
18	21	47.7	9	18	AAW41649	Peptide used in de
19	21	47.7	9	20	AAY47531	Immunogenic peptid
20	21	47.7	9	21	AAB10018	H. pylori beta ure
21	21	47.7	9	22	AAB86066	H. pylori beta ure
22	21	47.7	9	22	AAB36098	H. pylori beta ure
23	20	45.5	7	16	AAR72725	Mammalian ribosoma
24	20	45.5	7	18	AAW10781	Ferritin motif #.4
25	20	45.5	7	18	AAW17691	Substrate #1 for c
26	20	45.5	7	21	AAY83856	Ribonucleotide red
27	20	45.5	7	22	AAM40196	HLL binding site c
28	20	45.5	7	22	AAM46677	HLL binding site c
29	20	45.5	8	14	AAR82118	Heavy chain VHH-1
30	20	45.5	9	19	AAW87445	Peptide determinat
31	20	45.5	9	22	AAB56069	Vascular dementia
32	20	45.5	9	22	AAU28452	DPI cryptic digest
33	20	45.5	9	22	AAU26098	Depression Associa
34	20	45.5	9	22	AAB89808	HIV apl20 protein
35	20	45.5	9	22	AAB89809	HIV apl20 protein
36	20	45.5	9	22	AAB89810	HIV apl20 protein
37	20	45.5	9	22	AAB89811	HIV apl20 protein
38	19	43.2	5	16	AAR77436	Human apolipoprote
39	19	43.2	6	14	AAR29780	Selective factor 1
40	19	43.2	6	19	AAW41898	Peptide used in ra
41	19	43.2	7	16	AAR78178	Dog alpha IMI inte
42	19	43.2	7	16	AAW53654	Canine alpha IMI i
43	19	43.2	7	19	AAW65100	Canine beta interi
44	19	43.2	7	20	AAW57500	Canine alphaIM int
45	19	43.2	7	20	AAW93841	Fragment of canine

ALIGNMENTS

RESULT 1
AAR95071
ID AAR95071 standard; peptide; 9 AA.
XX
AC AAR95071;
XX
DT 17-DEC-1996 (first entry)
XX
DE Human cartilage glycoprotein 39 (HC gp 39) fragment (1).
XX
KW Human; cartilage; glycoprotein 39; HC gp 39; autoantigen;
KW induction; T-cell; tolerance; mediated; destruction;
KW rheumatoid arthritis; disease model; diagnosis.
XX
OS Homo sapiens.
XX
PN WO9613517-A1.
XX
PD 09-MAY-1996.
XX
PF 25-OCT-1995; 95WO-EP04201.
XX
PR 07-APR-1995; 95EP-0200886.
PR 27-OCT-1994; 94EP-0203128.
XX
PA (AKZO) AKZO NOBEL NV.
XX
PI Boots AMH, Verheijden GEM;
XX
DR WPI; 1996-239447/24.
XX
PT Peptide(s) derived from human cartilage gp 39 autoantigen - used
PT to induce specific T cell tolerance esp. for treatment of
PT rheumatoid arthritis, also useful in producing arthritic animal

```
PC   a Beta
XX
PA   Patent No. WO9747314-A1; Page 28; 47pp; English.
XX
CC   The present peptide, derived from the human cartilage glycoprotein,
CC   HC gp 39 (or YKL-40) autoantigen, can be used to induce specific T cell
CC   tolerance to HC gp 39 in a patient with T cell mediated cartilage
CC   destruction, esp. rheumatoid arthritis (RA). It may also be used to
CC   induce RA in mice, etc. (human disease models), and to diagnose the
CC   presence of activated T cells autoreactive with HC gp 39. Eg. to
CC   identify patients suitable for tolerising treatment.
CC   Therapeutically the peptide is given in a dose of 0.1-1.0 mg/kg
CC   by injection, or 10-50 micrograms/kg in a preferred dosage mode. The
CC   peptide has a specific tolerising effect on autoreactive T cells,
CC   i.e. leaves other components of the immune system intact, and
CC   therefore should have few side effects. In an assay to determine
CC   the peptide's binding to HLA DR mols., specifically DR4Dw4 (known
CC   to be associated with an increased risk of RA development), the
CC   peptide gave a 3rd blot 1/2(50) (micromolar) value of 0.04 compared
CC   to 0.5 for the control peptide HA(307-319).
XX
SQ   Sequence    9 AA;


  Query Match             100.0%;  Score 44;  DB 1;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 3.4e-05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 FGRSFTLAS 9
           |||||||||
Db       1 fgrsftlas 9


RESULT    3
AAW57025
ID   AAW57025 standard; peptide; 8 AA.
XX
AC   AAW57025;
XX
DT   28-JUL-1998  (first entry)
XX
DE   Enzyme inhibitor peptide SEQ ID NO:226.
XX
KW   Enzyme inhibitor; t-PA; u-PA; chymotrypsin; serine protease; active;
KW   latent; substrate subtraction phage display peptide library;
KW   identification; kinase; phosphatase; serpin.
XX
OS   Homo sapiens.
XX
PN   WO9747314-A1.
XX
PD   18-DEC-1997.
XX
PF   10-JUN-1997;   97WO-US09760.
XX
PR   10-JUN-1996;   96US-0019495.
XX
PA   (SCRI-) SCRIPPS RES INST.
XX
PI   Ke S;  Madison EL;
XX
DR   WPI; 1998-062746/06.
XX
PT   Substrate subtraction phage display peptide libraries - used to
PT   distinguish between active and latent forms of enzyme, e.g. serine
PT   protease
XX
PS   Claim 25; Page 117; 138pp; English.
XX
CC   The present sequence represents an enzyme inhibitor peptide used in
CC   the method of the invention to distinguish between t-PA and u-PA. The
CC   present invention describes a substrate subtraction library for the
CC   identification of peptide substrates selective between a first enzyme
CC   (E1) and a second enzyme (E2), comprising a collection of different
CC   peptides, substantially lacking peptides that are effective substrates
CC   for E1. Also described are: (1) a method (M1) for identifying peptide
CC   substrates selective between a first enzyme (E1) and a second enzyme
CC   (E2); (2) a compound comprising the amino acid sequence of a peptide
CC   identified by M1; (3) a polypeptide for use as an enzyme inhibitor
CC   comprising one of 237 amino acid sequences (see AAW56861 to AAW56947,
CC   and AAW56949 to AAW57038); (4) a recombinant DNA vector comprising DNA
CC   (1) encoding a protease inhibitor including the sequence identified by
CC   the M1; (5) a prokaryotic or eukaryotic cell containing the vector of
CC   (4); (6) an antibody (Ab) immunoreactive with at least one of the
CC   peptides identified by M1; and (7) a diagnostic assay for distinguishing
```

CC between active and latent forms of protease inhibitors, that uses (Ab).
CC The library and method are used for distinguishing between active and
CC latent forms of enzyme inhibitors, e.g. proteases, kinases and
CC phosphatases. (Ab) are used for affinity purification of recombinant
CC peptides and in the identification of naturally occurring protease
CC inhibitors. Enzyme- inhibiting peptides identified can be used to treat a
CC serpin deficiency or a disorder of serine proteases.
XX
SQ Sequence 8 AA;

Query Match 52.8%; Score 23; DB 1; Length 8;
Best Local Similarity 80.0%; Pred. No. 6.4e-05;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 GRSFI 6
|:|||
Db 2 GrsYI 7

RESULT 4
AAR95072
ID AAR95072 standard; peptide; 9 AA.
XX
AC AAR95072;
XX
DT 17-DEC-1996 (first entry)
XX
DE Human cartilage glycoprotein 39 autoantigen derived peptide (11).
XX
KW Human; cartilage; glycoprotein 39; hC gp-39; autoantigen;
KW induction; T-cell; tolerance; mediated; destruction;
KW rheumatoid arthritis; disease model; diagnosis.
XX
OS Homo sapiens.
XX
PN WO9613517-A1.
XX
PD 09-MAY-1996.
XX
PF 25-OCT-1995; 95WO-EP04201.
XX
PR 27-OCT-1994; 94EP-0203128.
XX
PA (AKZO) AKZO NOBEL NV.
XX
PI Boots AMH, Verheijden GPM;
XX
DR WPI; 1996-239447/24.
XX
PT Peptide(s) derived from human cartilage gp-39 autoantigen used
PT to induce specific T cell tolerance, esp. for treatment of
PT rheumatoid arthritis, also useful in producing arthritic animal
PT models
XX
PS Claim 1; Page 28; 37pp; English.
XX
CC The present peptide, derived from the human cartilage glycoprotein
CC 39 (hC gp-39) autoantigen, can be used to induce specific T cell
CC tolerance to hC gp-39 in a patient with T-cell mediated cartilage
CC destruction, esp. rheumatoid arthritis (RA). It may also be used to
CC induce RA in mice, etc. (human disease models), and to diagnose the
CC presence of activated T cells autoreactive with hC gp-39, i.e. to
CC identify patients suitable for tolerising treatment.
CC Therapeutically the peptide is given in a dose of 0.1-10, preferably
CC by injection, or 10-50 microg/kg to produce a disease model. The
CC peptide has a specific tolerising effect on autoreactive T cells,
CC but leaves other components of the immune system intact, and
CC therefore should have few side effects. In an assay to determine
CC the peptide's binding to HLA-DP mols., specifically DP4(w4) (known
CC to be associated with an increased risk of RA development), the
CC peptide gave a DOT blot IC(50) (micromolar) value of 0.08 compared
CC to 0.55 for the control peptide IBA(w2-w16)F.
XX
SQ Sequence 9 AA;

Query Match 52.8%; Score 23; DB 1; Length 9;
Best Local Similarity 100.0%; Pred. No. 6.4e-05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 FTLAS 9
|||||
Db 1 FTLAS 5

RESULT 5
AAY69227
ID AAY69227 standard; peptide; 9 AA.
XX
AC AAY69227;
XX
DT 30-MAY-2000 (first entry)
XX
DE Peptide fragment of human cartilage gp-39 (HC gp-39) protein.
XX
KW Human cartilage gp-39 protein; HC gp-39 protein; T cell activator;
KW inflammatory disease; lymphocyte; antigen; bystander suppression;
KW antigen-specific T-cell tolerance; autoimmune disorder;
KW inflammatory disease; rheumatoid arthritis;
KW autoimmune inflammatory disease; autoantigen.
XX
OS Homo sapiens.
XX
PN WO200004917-A2.
XX
PD 03-FEB-2000.
XX
PF 19-JUL-1999; 99WO-EP05331.
XX
PR 23-JUL-1998; 98EP-0202471.
XX
PA (AKZO) AKZO NOBEL NV.
XX
PI Miltenburg AMM, Boots AMH;
XX
DR WPI; 2000-182537/16.
XX
PT Use of cartilage protein or peptides for treating inflammatory diseases
PT especially rheumatoid arthritis
XX
PS Claim 2; Page 17; 29pp; English.
XX
CC AAY69226-33 represent fragments of the human cartilage gp-39 (HC gp-39)
CC protein. The peptides are T cell activators. The peptides are used for
CC preventing inflammatory diseases, by modulating lymphocytes reactive
CC to antigens other than HC gp-39, present in the same tissue of
CC HC gp-39. By the induction of antigen specific T cell tolerance,
CC autoimmune disorders can be treated by bystander suppression. The
CC HC gp-39 peptides are useful for preventing inflammatory diseases
CC especially rheumatoid arthritis. The fragments may also be used for
CC treating autoimmune inflammatory diseases. The HC gp-39 peptides are
CC especially beneficial in autoimmune conditions in which the
CC autoantigens associated with the condition are unknown
XX
SQ Sequence 9 AA;

Query Match 52.8%; Score 23; DB 1; Length 9;
Best Local Similarity 100.0%; Pred. No. 6.4e-05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 FTLAS 9
|||||
Db 1 FTLAS 5

```
RESULT   6
AAY[illegible]
ID   AAY[illegible] standard; peptide; 9 AA.
XX
AC   AAY[illegible];
XX
DT   [illegible] MAY [illegible] (first entry)
XX
DE   Human 54 kDa cartilage protein YKL 54, residues 268-276.
XX
KW   Human 54 kDa cartilage protein; YKL 54; epitope; autoreactive T cell;
KW   [illegible]tion; desensitisation; autoantigen; rheumatoid arthritis;
KW   HC gp 39; human cartilage glycoprotein 39; YKL 40; bystander suppression;
KW   [illegible] disorder; Graves' disease; primary glomerulonephritis;
KW   inflammatory bowel disease; diagnosis.
XX
OS   Homo sapiens.
XX
PN   WO[illegible]24 A2.
XX
PD   [illegible] 2000.
XX
PF   [illegible] 1999;   99WO EP05050.
XX
PR   23 JUL 1998;   98EP 0202470.
XX
PA   (ARLA) , AKZO N-BEL NV.
XX
PI   Vervoordeldonk MJB, Boots AMH;
XX
DR   WPI; [illegible]-[illegible]44/29.
XX
PT   New [illegible] epitope peptides useful for inducing specific T cell
PT   [illegible] from autoantigen for treating autoimmune disorders,
PT   especially rheumatoid arthritis
XX
PS   Claim 1; Page 36; 26pp; English
XX
CC   Sequences AAY[illegible] Y[illegible] respectively represent amino acids 268-276 and
CC   amino acids 263-276 of human chondrocyte protein YKL 54. The peptides
CC   may be used in pharmaceutical compositions for the induction of specific
CC   T cell tolerance (desensitisation) to an autoantigen in patients
CC   suffering from autoimmune disorders, especially rheumatoid arthritis. A
CC   frequently found autoantigen in rheumatoid arthritis sufferers is human
CC   cartilage glycoprotein 39 (HC gp 39, also known as YKL 40). Several
CC   epitopes of HC gp 39 are recognised by peripheral blood T cells from
CC   rheumatoid arthritis patients; one such epitope, HC gp 39 residues
CC   [illegible] was recognised by T cells from 9 out of 10 patients tested.
CC   YKL 54 shares significant sequence identity to the gp 39; YKL 54 residues
CC   263-276 (AAY[illegible]) constitutes a mimicry epitope of HC gp 39 (263-275),
CC   and may be used for desensitisation of T cells with reactivity to HC
CC   gp 39 (263-275) in rheumatoid arthritis patients. Rheumatoid arthritis
CC   may be treated by the systemic administration of compositions comprising
CC   the peptides of the invention, optionally in combination with existing
CC   treatments. The peptides may also be used to modulate lymphocytes
CC   that are reactive to autoantigens other than YKL 54 and HC gp 39 but
CC   which are present in the same tissue as those autoantigens. By the
CC   activity of antigen specific T cells to remove, autoimmune disorders can
CC   be treated by bystander suppression. The conditions that may be treated
CC   via this approach include Graves' disease, primary glomerulonephritis
CC   and inflammatory bowel disease. The peptides, alone or in a detection
CC   kit, may additionally be used in a diagnostic capacity. The peptides
CC   of the invention have a specific effect on autoreactive T cells whilst
CC   leaving other components of the immune system intact. Therefore, the
CC   peptides should be safe, having no side effects.
XX
SQ   Sequence   9 AA;

  Query Match             92.9%;  Score 26;  DB 21;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 9.4e-05;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       5 FTLAS 9
           |||||
Db       1 ttlas 5


RESULT   7
AAW92347
ID   AAW92347 standard; peptide; 7 AA.
XX
AC   AAW92347;
XX
DT   09-APR-1999  (first entry)
XX
DE   Latent fluorescent peptide #11.
XX
KW   Fluorescent peptide; masking group; fluorescent group; sensitive; stable;
KW   phosphatase activity; phosphoric acid.
XX
OS   Synthetic.
XX
PN   JP11012297-A.
XX
PD   19-JAN-1999.
XX
PF   21-APR-1998;   98JP-0126684.
XX
PR   30-APR-1997;   97JP-0126463.
XX
PA   (NISH/) NISHIKATA S.
XX
DR   WPI; 1999-148570/13.
XX
PT   New latent fluorescent peptide(s) - useful for determining protein
PT   phosphatase activity
XX
PS   Claim 5; Page 10; 11pp; Japanese.
XX
CC   AAW92347-AAW92368 are novel latent fluorescent (optionally protected)
CC   peptides having a self fluorescent group and masking group in the
CC   molecule and phosphoric acid groups in the masking group and between the
CC   fluorescent and masking groups. The invention also describes peptide
CC   derivatives of formula (A1-X-A2), especially composed of 3-42 amino
CC   acids; X = amino acid residue comprising phosphoric acid residue; A1,
CC   A2 = (optionally protected) amino acid residue, (optionally protected)
CC   peptide chain, one comprises a masking group at the terminal or side
CC   chain and the other has a self fluorescent group capable of being masked
CC   with the intramolecular masking group. The peptides are useful for
CC   determining protein phosphatase activity by the changes of intensity of
CC   fluorescence. The fluorescent peptides are stable reagents for sensitive
CC   and rapid determination of protein phosphatase activity in crude sample
CC   even contaminated sample containing phosphoric acid ion.
XX
SQ   Sequence   7 AA;

  Query Match             50.0%;  Score 22;  DB 20;  Length 7;
  Best Local Similarity   80.0%;  Pred. No. 6.4e+05;
  Matches    4;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       2 GSSFI 6
           | ||
Db       1 grkit 5


RESULT   8
AAG63003
ID   AAG63003 standard; peptide; 7 AA.
XX
AC   AAG63003;
XX
DT   01-OCT-2001  (first entry)
```

XX
DE Complementarity determining region 3 (CDR3) of VH chain of crone Gli...
XX
KW Antibody; light chain; VL; amyloid protein; blood brain barrier;
KW endothelial cell; brain cell antigen; inflammation; adhesion molecule;
KW transferrin receptor; neurological disease; Alzheimer's disease;
KW prion disease; AIDS-related dementia; epilepsy; brain injury.
XX
OS Homo sapiens.
XX
PN W[illegible] A2.
XX
PD 21-[illegible]-[illegible].
XX
PF 27-NOV-2000; 2000WO-GB04501.
XX
PR 18-DEC-1999; 99US-0170599.
XX
PA (CAMB-) CAMBRIDGE ANTIBODY TECHNOLOGY.
XX
PI Webster C, Osbourn J, Ward G, Miller K;
XX
DR WPI; 2001-398131/42.
XX
PT Mixture or panel of antibodies for selecting specific binding members
PT that cross the blood brain barrier, for use in delivering different
PT molecules and treating neurological diseases
XX
PS Claim 1; Page 76; 109pp; English.
XX
CC AAG62-70-AAG630-5 represent complementarity determining region 3 (CDR3)
CC of VL and VH chains of antibodies of the invention. The specification
CC describes a mixture or panel of 5 different specific binding members,
CC each comprising an antibody VH and/or VL variable domain and capable,
CC when displayed on the surface of filamentous bacteriophage particles or
CC in the case of a specific binding member comprising the VH and/or VL
CC variable domain when bound to human serum amyloid protein, to pass
CC through a mammalian blood brain barrier (BBB). The panel is useful for
CC the selection of specific binding members with a desired property such
CC as ability to cross BBB, ability to bind endothelial cells or other brain
CC cell antigen, ability to bind areas of inflammation in the brain or BBB
CC breakdown or ability to bind intracellular adhesion molecules and to bind
CC transferrin receptor. The antibodies are useful in diagnosis, prophylaxis
CC and treatment of human or animal body, including neurological diseases,
CC such as Alzheimer's disease, prion disease, AIDS-related dementia,
CC epilepsy and traumatic brain injury and any diseases involving
CC inflammation occurring within the brain or central nervous system.
XX
SQ Sequence 7 AA;

Query Match 50.0%; Score 22; DB 22; Length 7;
Best Local Similarity 66.7%; Pred. No. 6.4e-05;
Matches 4; Conservative 1; Mismatches 1; Indels 1; Gaps 0;

Qy 2 GRSFT 6
|||:
Db 1 grsft 5

RESULT 9
AAY10406
ID AAY10406 standard; Peptide; 9 AA.
XX
AC AAY10406;
XX
DT 12-MAY-[illegible] (first entry)
XX
DE T cell epitope/MHC ligand SEQ ID NO:356.
XX
KW Cytotoxic T lymphocyte response; CTL; antigen; lymphatic system;
KW immunisation; tumour; infectious disease; immunotherapy; cancer;
KW malignant melanoma; viral disease; hepatitis; AIDS.
XX
OS Synthetic.
OS Rotavirus.
XX
PN WO9902183-A2.
XX
PD 21-JAN-1999.
XX
PF 10-JUL-1998; 98WO-US14289.
XX
PR 10-DEC-1997; 97US-0988820.
PR 10-JUL-1997; 97CA-2209815.
XX
PA (CTLI-) CTL IMMUNOTHERAPIES CORP.
XX
PI Kuendig TM, Simard JJL;
XX
DR WPI; 1999-120514/10.
XX
PT Inducing a cytotoxic T lymphocyte response by maintaining a level
PT of antigen in the lymphatic system of a mammal so as to provide a
PT sustained CTL response, used to treat, e.g. AIDS
XX
PS Disclosure; Page 37; 199pp; English.
XX
CC The present invention describes a method of inducing and/or sustaining
CC an immunological cytotoxic T lymphocyte (CTL) response in a mammal. The
CC method comprises: (a) delivering an antigen to the mammal at a level to
CC induce an immunological CTL response in the mammal; and (b) maintaining
CC the level of the antigen in the mammal's lymphatic system to maintain
CC the immunologic CTL response. The method can be used: for the delivery of
CC e.g. a differentiation antigen, a tumour-specific multi-lineage antigen,
CC an embryonic antigen, an oncogene antigen, a mutated tumour suppressor
CC gene antigen, or a viral antigen. They can be used for the treatment of
CC disease such as cancer, e.g. malignant melanoma or infectious disease,
CC e.g. viral disease such as hepatitis or AIDS. Sustained antigen delivery
CC to the lymphatic system provides for potent CTL stimulation that takes
CC place in the milieu of the lymphoid organ, and it sustains stimulation
CC that is necessary to keep CTL active, cytotoxic and recirculating
CC through the body. AAY10071 to AAY10699 represent examples of peptide
CC antigens given in the present invention.
XX
SQ Sequence 9 AA;

Query Match 50.0%; Score 22; DB 20; Length 9;
Best Local Similarity 80.0%; Pred. No. 6.4e-05;
Matches 4; Conservative 1; Mismatches 0; Indels 1; Gaps 1;

Qy 2 GRSFT 6
|||||
Db 3 qrsfs 7

RESULT 10
AAB34002
ID AAB34002 standard; Peptide; 9 AA.
XX
AC AAB34002;
XX
DT 02-FEB-2001 (first entry)
XX
DE Human secreted protein encoded by cDNA #41.
XX
KW Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
KW antiallergic; hepatotropic; antidiabetic; antiinflammatory; antiulcer;
KW vulnerary; anticonvulsant; antibacterial; antitumour; antiparasitic;
KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
KW neurological disease; infection; human; secreted protein.
XX
OS Homo sapiens.
XX
PN WO200056765-A1.

XX
PN [illegible]
XX
PD [illegible]
XX
PF [illegible]
XX
PR [illegible]
XX
PA [illegible]
XX
PI [illegible]
XX
DR [illegible]
XX
PT [illegible]
XX
PS [illegible]
XX
CC [illegible]
XX
SQ Sequence [illegible]

Query Match [illegible]; Score 22; DB 21; Length 9;
Best Local Similarity [illegible]; Pred. No. [illegible];
Matches [illegible]; Conservative [illegible]; Mismatches [illegible]; Indels [illegible]; Gaps [illegible];

Qy [illegible]
Db [illegible]

RESULT 11
AAW[illegible]
ID AAW[illegible] standard; peptide; 4 AA.
XX
AC AAW[illegible];
XX
DT [illegible] (first entry)
XX
DE [illegible]
XX
KW [illegible]
XX
OS [illegible]
XX
FH Key Location/Qualifiers
FT Modified-site [illegible]
FT /note= "Z Phe psi[PO2CH2] Gly, where Z is benzyloxy
FT carbonyl and psi[PO2CH2] indicates replacement of
FT the peptide linkage -CONH- between Phe and Gly by
FT the group -PO2CH2-"
XX

PN EP725075 A1.
XX
PD 07-AUG-1996.
XX
PF 02-FEB-1996; 96EP-0400229.
XX
PR 06-FEB-1995; 95FR-0001428.
XX
PA (COMS) COMMISSARIAT ENERGIE ATOMIQUE.
XX
PI Dive V, Jiracek J, Yiotakis A;
XX
DR WPI; 1996-356059/36.
XX
PT New peptide derivs. contg. phosphinic acid gp. replacing an amide
PT bond are highly specific inhibitors of endopeptidase 24.15, for
PT treating hypothermia, hypertension, cancer, Alzheimer's disease etc.
XX
PS Disclosure; Page 9; 18pp; French.
XX
CC The sequence is a specific example of new peptide derivatives containing
CC the sequence -Phe psi[PO2CH2] X-Y Z- in which Y Arg or Lys; X and Z'
CC natural or pseudo-amino acids (preferably X is Gly, Ala or Leu and Z' is
CC Met, Nle, Ala or Phe); and psi[PO2CH2] indicates replacement of the
CC peptide linkage -CONH- between Phe and X by the group -PO2CH2-.
CC These peptides are inhibitors of the zinc dependent endopeptidase
CC EC.3.4.24.15 and so prevent degradation of e.g. somatostatin, bradykinin,
CC angiotensin, neurotensin, substance P, dynorphin etc. and may prevent
CC maturation of ras oncoprotein. They are useful in treatment of pain,
CC hypothermia, arterial hypertension, cancer and Alzheimer's disease.
CC They are very selective for 24.15 with no significant action on other
CC zinc endopeptidases such as 24.16, and are more stable, chemically, than
CC phosphonamide peptide derivatives.
XX
SQ Sequence 4 AA;

Query Match 47.7%; Score 21; DB 17; Length 4;
Best Local Similarity 100.0%; Pred. No. 6.4e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FGRS 4
||||
Db 1 fgrs 4

RESULT 12
AAY33132
ID AAY33132 standard; Protein; 5 AA.
XX
AC AAY33132;
XX
DT 16-NOV-1999 (first entry)
XX
DE Human umbilical cord specific epitope from clone UCV P01* #2.
XX
KW Tissue specific epitope; umbilical cord; human; endothelial cell; vein;
KW tumor cell; tumor tissue; delivery agent; imaging; artery; radial; liver;
KW coronary; mammary; saphenal; femoral; placenta; kidney; heart;
KW central nervous system.
XX
OS Homo sapiens.
XX
PN WO9945020 A1.
XX
PD 10-SEP-1999.
XX
PF 04-MAR-1999; 99WO-US004691.
XX
PR 04-MAR-1998; 98WO-US004188.
PR 10-SEP-1998; 98US-0154404.
XX
PA (BIOT) BIOTECHNOLOGY GENERAL CORP.

```
XX
PI   Vogel T, Peter A;
XX
PN   WO: [illegible] A1/45.
XX
PT   New peptides useful for directing therapeutics to cancer tissues and
PT   endothelial cells
XX
PS   Example 1; Page 68; 70pp; English.
XX
CC   This invention describes novel non-naturally occurring pharmaceutically
CC   active peptide epitopes (I) comprising the tripeptide, G..Gly Arg. (I)
CC   are useful for binding to endothelial cell and tumor cells/tissues, and
CC   are useful as delivery agents. The products described in the invention
CC   are also useful for imaging an organ, preferably an artery (umbilical
CC   cord artery, radial artery, coronary artery or mammary artery), a vein
CC   (umbilical cord vein, saphenal vein or femoral vein), placenta, tumor
CC   tissue, kidney, heart or liver. The artery is preferably damaged, and is
CC   a coronary artery. They are also useful for treating an organ in vivo,
CC   preferably an artery (umbilical cord artery, radial artery, coronary
CC   artery or a mammary artery, preferably a damaged coronary artery), vein
CC   (umbilical cord vein, saphenal vein or femoral vein), placenta, tumor
CC   tissue, kidney, heart, liver, or central nervous system. AAY65254-Y65252
CC   represent epitopes described in the method of the invention.
XX
SQ   Sequence   5 AA;


  Query Match             47.7%;  Score 21;  DB 20;  Length 5;
  Best Local Similarity   100.0%;  Pred. No. 6.4e-05;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 RGSE 4
             ||||
Db       1 RGSE 4


RESULT  13
AAY62476
ID   AAY62476 standard; peptide; 5 AA.
XX
AC   AAY62476;
XX
DT   22-MAY 2000  (first entry)
XX
DE   S. clava antimicrobial styelin A peptide fragment (residues 4-7).
XX
KW   Antimicrobial; styelin; microbial degradation; bactericide; viricide;
KW   fungicide; plant protection; styelin A.
XX
OS   Styela clava.
XX
PN   WO9822587-A1.
XX
PD   28-MAY 1998.
XX
PF   14 APR 1998;   98WO-US007430.
XX
PR   24-DEC 1997;   97US-0068802.
PR   23-JAN 1998;   98US-0072885.
PR   10-FEB 1998;   98US-0075026.
XX
PA   (REGE ) UNIV CALIFORNIA.
XX
PI   Zhao C;  Lee I;  Lehrer RI;
XX
DR   WPI; 2000-[illegible]480/15.
XX
PT   Antimicrobial styelin peptides isolated from Styela clava useful for
PT   preserving materials vulnerable to microbial degradation, and for
PT   protecting plants and animals against pathogenic microbes
XX
PS   Example 3; Page 22; 45pp; English.
XX
CC   The invention relates to antimicrobial peptides (styelins) (and
CC   acetylated forms, salts, amides and esters of styelins) isolated from
CC   Styela clava and characterized by specific patterns of basic and
CC   hydrophobic amino acid side chains which display a broad spectrum of
CC   antimicrobial activities. The peptides are of a specified formula and
CC   display a wide range of antimicrobial activities and are therefore useful
CC   for preserving materials susceptible to microbial degradation, for
CC   protecting plants against bacterial infection and in the therapeutic and
CC   prophylactic protection of animals against bacteria, fungi and viruses.
CC   The compounds may also be used as standards in antimicrobial assays and
CC   as affinity ligands for absorption of counterpart structures in microbes,
CC   including viruses. The present sequence represents a partial fragment of
CC   styelin A, used for designing primer sequences.
XX
SQ   Sequence   5 AA;


  Query Match             47.7%;  Score 21;  DB 21;  Length 5;
  Best Local Similarity   60.0%;  Pred. No. 6.4e-05;
  Matches    3;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 FGRSE 5
             |:::|
Db       1 tgkat 5


RESULT  14
AAY16509
ID   AAY16509 standard; peptide; 8 AA.
XX
AC   AAY16509;
XX
DT   06-AUG 1999  (first entry)
XX
DE   Betide sequence of the invention.
XX
KW   Betide; betide library; chain elongation; betidamino acid;
KW   screening; cyclic.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Misc difference 1..8
FT                   /note  "the betide is cyclised via these residues"
FT   Modified-site   4
FT                   /note  "(beta-methyl gamma 2 naphthyl) : of
FT                          L-form (amido)Gly"
XX
PN   WO9618642-A1.
XX
PD   20-JUN-1996.
XX
PF   15-DEC-1995;   95WO-US16205.
XX
PR   16-DEC-1994;   94US-0358184.
XX
PA   (SALK ) SALK INST BIOLOGICAL STUDIES.
XX
PI   Porter JS,  Rivier JEF;
XX
DR   WPI; 1996-300570/30.
XX
PT   New peptide(s) containing betidamino acid residues  used in screening
PT   for novel activities, for increased biopotencies or for other predd.
PT   properties
XX
PS   Example 7; Page 54; 77pp; English.
XX
CC   The specification describes a method for use of a betide library
CC   using a chain elongation protocol. A betide is a peptide which
CC   contains a betidamino acid, which is a non-acylated version of
CC   aminoglycine, alpha aminoalanine, alpha aminosarcosine
```

```
XX   His acetylated diaminobutyrine or his methylated diaminoalanine
XX   where the N  or alpha amino site is destined to become part of
XX   the backbone of the peptide chain of residues, and the N  or
XX   beta amino site is substituted with an acyl group that represents a
XX   lipid or alpha amino acid side chain, so that the betidamino acid
XX   represents either the alpha amino acid. The libraries can be used
XX   for screening for novel activities, for increased bioavailability or
XX   other preferred properties. They can also be used to provide
XX   information for designing novel peptides with improved properties.
XX   The present sequence represents a betide of the invention.
XX
SQ   Sequence   8 AA;

  Query Match          47.7%;  Score 21;  DB 2;  Length 8;
  Best Local Similarity   80.0%;  Pred. No. 6.4e-05;
  Matches    4;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 GRSFI 6
              |: ||
Db          2 GKTFI 6


RESULT 2
AAW43728
ID   AAW43728 standard; peptide; 8 AA.
XX
AC   AAW43728;
XX
DT   25-FEB-1998  (first entry)
XX
DE   Marker protein C-terminal fragment.
XX
KW   Marker protein; target protein; molecular weight; density determination.
XX
OS   Synthetic.
XX
PN   EP802254-A2.
XX
PD   22-MAR-1998.
XX
PF   20-AUG-1997;   97EP-0402024.
XX
PR   20-AUG-1996;   96JP-0246859.
XX
PA   (SEIK ) SEIKO INSTR INC.
XX
PI   Fujita Y;  Matsuzaki Y;
XX
DR   WPI; 1998-184084/16.
XX
PT   Determination of molecular weight and/or amount of target protein
PT   by electrophoresis using marker consisting of proteins that have
PT   different molecular weights and are present in different amounts
XX
PS   Claim; Sure; Page 5; 11pp; English.
XX
CC   This peptide is used in a novel method for determining the molecular
CC   weight and/or amount of a target protein. The method involves
CC   subjecting the target protein to electrophoresis simultaneously
CC   with a marker consisting of proteins that have different molecular
CC   weights and are present in different amounts, staining the target
CC   protein and the marker proteins, and comparing the resulting bands by
CC   eye or with a densitometer. The molecular weight can be determined by
CC   measuring the position of the target protein band with the various
CC   marker proteins. The amount can be determined
CC   by measuring the density of the target protein band with the various
CC   marker proteins bands.
XX
SQ   Sequence   8 AA;

  Query Match          47.7%;  Score 21;  DB 2;  Length 8;
  Best Local Similarity   80.0%;  Pred. No. 6.4e-05;
  Matches    4;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 GRSFI 6
              |: ||
Db          2 GKTFI 6


Search completed: August 26, 2002, 06:27:08
Job time: 989 sec
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 26, 2002, 06:10:39 ; Search time 5.76 Seconds
(without alignments)
4.345 Million cell updates/sec

Title: US-09-744-282-1
Perfect score: 44
Sequence: 1 FQRSFILAS 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 6.7 %

Minimum DB seq length: 0
Maximum DB seq length: 9

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/ina/5_COMB_pep.*
2: /cgn2_6/ptodata/1/ina/6_COMB_pep.*
3: /cgn2_6/ptodata/1/ina/6A_COMB.pep.*
4: /cgn2_6/ptodata/1/ina/PCTUS_COMB.pep.*
5: /cgn2_6/ptodata/1/ina/backfiles1.pep.*
6: /cgn2_6/ptodata/1/ina/backfiles2.pep.*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match %	Score	Match Length	DB	ID	Description
1	44	100.0	9	1	US-08-619-645-1	Sequence 1, Appli
2	44	100.0	9	2	US-08-641-197-1	Sequence 1, Appli
3	44	100.0	9	1	US-08-619-645-2	Sequence 2, Appli
4	44	100.0	9	2	US-08-634-493-2	Sequence 2, Appli
5	44	100.0	7	2	US-09-070-756-12	Sequence 12, Appl
6	44	100.0	7	2	US-09-070-756-12	Sequence 12, Appl
7	44	47.7	8	1	US-08-409-809-1	Sequence 1, Appli
8	44	47.7	8	2	US-08-450-260A-169	Sequence 169, App
9	44	47.7	9	1	US-08-215-865A-54	Sequence 54, Appl
10	44	45.5	7	1	US-08-136-743B-53	Sequence 53, Appl
11	44	45.5	7	1	US-08-213-867A-5	Sequence 5, Appli
12	44	45.5	7	4	US-09-040-216-26	Sequence 26, Appl
13	44	45.5	9	1	US-08-446-333-35	Sequence 35, Appl
14	44	45.5	9	5	PCT-US91-07506-35	Sequence 35, Appl
15	19	43.2	5	1	US-08-456-840-32	Sequence 32, Appl
16	19	43.2	5	1	US-08-266-407A-32	Sequence 32, Appl
17	19	43.2	5	2	US-08-892-544-32	Sequence 32, Appl
18	19	43.2	6	1	US-08-049-786-2	Sequence 2, Appli
19	19	43.2	6	1	US-08-238-329-2	Sequence 2, Appli
20	19	43.2	6	1	US-08-290-448A-41	Sequence 41, Appl
21	19	43.2	6	1	US-08-290-448A-41	Sequence 41, Appl
22	19	43.2	6	1	US-08-175-069A-41	Sequence 41, Appl
23	19	43.2	6	4	US-08-461-969B-41	Sequence 41, Appl
24	19	43.2	6	4	US-08-464-000-41	Sequence 41, Appl
25	19	43.2	7	1	US-08-173-497-22	Sequence 22, Appl
26	19	43.2	7	1	US-08-286-889-22	Sequence 22, Appl
27	19	43.2	7	1	US-08-485-618-22	Sequence 22, Appl
28	19	43.2	7	1	US-08-362-65[illegible]-22	Sequence 22, Appl
29	19	43.2	7	2	US-08-605-67[illegible]-22	Sequence 22, Appl
30	19	43.2	7	2	US-08-481-29[illegible]-22	Sequence 22, Appl
31	19	43.2	7	2	US-08-963-35[illegible]-22	Sequence 22, Appl
32	19	43.2	7	2	US-08-350-26[illegible]A-422	Sequence 422, App
33	19	43.2	7	2	US-08-751-70[illegible]A-44	Sequence 44, Appl
34	19	43.2	7	4	US-09-196-34[illegible]-22	Sequence 22, Appl
35	19	43.2	7	4	US-09-257-06[illegible]-5	Sequence 5, Appli
36	19	43.2	8	1	US-08-136-743B-44	Sequence 44, Appl
37	19	43.2	8	1	US-08-287-95[illegible]-22	Sequence 22, Appl
38	19	43.2	8	1	US-08-394-944-A-15	Sequence 15, Appl
39	19	43.2	8	2	US-08-350-260A-471	Sequence 471, App
40	19	43.2	8	4	US-09-040-216-14	Sequence 14, Appl
41	19	43.2	8	4	US-09-040-216-15	Sequence 15, Appl
42	19	43.2	8	5	PCT-US96-09512-18	Sequence 18, Appl
43	19	43.2	9	1	US-08-136-743B-43	Sequence 43, Appl
44	19	43.2	9	1	US-08-475-520-6	Sequence 6, Appli
45	19	43.2	9	4	US-09-04[illegible]-216-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1
US-08-619-645-1
; Sequence 1, Application US/08619645
; Patent No. 5736507
; GENERAL INFORMATION:
; APPLICANT: Boots, Anna M.H.
; APPLICANT: Verheijden, Gijsbertus F.M.
; TITLE OF INVENTION: No. 5736507 peptides derived from
; TITLE OF INVENTION: autoantigen for use in immunotherapy of autoimmune
; TITLE OF INVENTION: disease
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Akzo Nobel Patent Department
; STREET: 1300 Piccard Drive, Suite 209
; CITY: Rockville
; STATE: Maryland
; COUNTRY: U.S.A.
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25(EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/619,645
; FILING DATE: 25-MAR-1996
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Gormley, Mary E.
; REGISTRATION NUMBER: 34,400
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 947-4455
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-619-645-1

Query Match 100.0%; Score 44; DB 1; Length 9;
Best Local Similarity 100.0%; Pred. No. 2.5e-07;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FQRSFILAS 9
|||||||||
Db 1 FQRSFILAS 9

```
RESULT   2
US-08-634-493-1
; Sequence 1, Application US/08634493
; Patent No. 5843449
; GENERAL INFORMATION:
;    APPLICANT:  A.M.H. Boots
;    APPLICANT:  G.F.M. Verheijden
;    APPLICANT:  E.S. Bos
;    TITLE OF INVENTION:  No. 5843449el Peptides derived from autoantigen for use
;    TITLE OF INVENTION:  in Immunotherapy of Autoimmune Diseases
;    NUMBER OF SEQUENCES:  10
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Akzo No. 5843449el Patent Department
;      STREET:  1300 Piccard Drive, Suite 206
;      CITY:  Rockville
;      STATE:  Maryland
;      COUNTRY:  U.S.A.
;      ZIP:  20850
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/634,493
;      FILING DATE:  18 APR 1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/619,645
;      FILING DATE:  25 MAR 1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/EP95/04201
;      FILING DATE:  25 OCT 1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  NL 9420 1287
;      FILING DATE:  27 OCT 1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  NL 9500 0865
;      FILING DATE:  07 APR 1995
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (301) 258 5200
;      TELEFAX:  (301) 977 0847
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  9 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-634-493-2

  Query Match             100.0%;  Score 44;  DB 2;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 2.7e-05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        1 FTLAS 5
              |||||
Db        1 FTLAS 5


RESULT   3
US-08-634-493-3
; Sequence 2, Application US/08634493
; Patent No. 5843449
; GENERAL INFORMATION:
;    APPLICANT:  Bos, E.S.; Amna M.H.
;    APPLICANT:  Verheijden, G.; Sjoerdsma E.M.
;    TITLE OF INVENTION:  No. 5736507el peptides derived from
;    TITLE OF INVENTION:  autoantigen for use in Immunotherapy of Autoimmune
;    TITLE OF INVENTION:  diseases
;    NUMBER OF SEQUENCES:  9
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Akzo No. 5736507el Patent Department
;      STREET:  1300 Piccard Drive, Suite 206
;      CITY:  Rockville
;      STATE:  Maryland
;      COUNTRY:  U.S.A.
;      ZIP:  20850
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25(EPO)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/619,645
;      FILING DATE:  25 MAR-1996
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Gormley, Mary E.
;      REGISTRATION NUMBER:  34,409
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (301) 947 4433
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  9 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-619-645-2


  Query Match             52.3%;  Score 23;  DB 1;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 1.7e-05;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        5 FTLAS 9
              |||||
Db        1 FTLAS 5


RESULT   4
US-08-634-493-2
; Sequence 2, Application US/08634493
; Patent No. 5843449
; GENERAL INFORMATION:
;    APPLICANT:  A.M.H. Boots
;    APPLICANT:  G.F.M. Verheijden
;    APPLICANT:  E.S. Bos
;    TITLE OF INVENTION:  No. 5843449el Peptides derived from autoantigen for use
;    TITLE OF INVENTION:  in Immunotherapy of Autoimmune Diseases
;    NUMBER OF SEQUENCES:  10
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Akzo No. 5843449el Patent Department
;      STREET:  1300 Piccard Drive, Suite 206
;      CITY:  Rockville
;      STATE:  Maryland
;      COUNTRY:  USA
;      ZIP:  20850
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC DOS/MS DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/634,493
;      FILING DATE:  18 APR-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/619,645
;      FILING DATE:  25 MAR-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/EP95/04201
;      FILING DATE:  25 OCT-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  NL 9420 1287
```

```
;       FILING DATE:  27-OCT-1994
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  NL 952008860
;       FILING DATE:  07-APR-1995
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (301) 258-5200
;       TELEFAX:  (301) 977-0847
;   INFORMATION FOR SEQ ID NO:  2:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  9 amino acids
;       TYPE:  amino acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  peptide
US-08-634,494-2


  Query Match             52.3%;  Score 23;  DB 2;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 1.7e+05;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       5 FTLAS 9
            |||||
Db       1 FTLAS 5


RESULT   5
US-09-070-756-11
; Sequence 11, Application US/09070756
; Patent No. 5917012
;  GENERAL INFORMATION:
;     APPLICANT:  Nishikata, Makoto
;     TITLE OF INVENTION:  NOVEL PEPTIDE DERIVATIVES
;     NUMBER OF SEQUENCES:  14
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  DAVID S. RESNICK; DIKE, BRONSTEIN, ROBERTS &
;       ADDRESSEE:  CUSHMAN
;       STREET:  130 Water Street
;       CITY:  Boston
;       STATE:  Massachusetts
;       COUNTRY:  US
;       ZIP:  02109
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/09/070,756
;       FILING DATE:  30 April 1998
;       CLASSIFICATION:  530
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Resnick, David S.
;       REGISTRATION NUMBER:  34,235
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (617)523-3400
;       TELEFAX:  (617)523-6440
;   INFORMATION FOR SEQ ID NO:  11:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  7 amino acids
;       TYPE:  amino acid
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  peptide
US-09-070-756-11


  Query Match             50.0%;  Score 22;  DB 2;  Length 7;
  Best Local Similarity   80.0%;  Pred. No. 1.7e+05;
  Matches    4;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       2 GRSFT 6
            || ||
Db       1 GREFT 5


RESULT   6
US-09-070-756-12
; Sequence 12, Application US/09070756
; Patent No. 5917012
;  GENERAL INFORMATION:
;     APPLICANT:  Nishikata, Makoto
;     TITLE OF INVENTION:  NOVEL PEPTIDE DERIVATIVES
;     NUMBER OF SEQUENCES:  14
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  DAVID S. RESNICK; DIKE, BRONSTEIN, R BERTS &
;       ADDRESSEE:  CUSHMAN
;       STREET:  130 Water Street
;       CITY:  Boston
;       STATE:  Massachusetts
;       COUNTRY:  US
;       ZIP:  02109
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/09/070,756
;       FILING DATE:  30 April 1998
;       CLASSIFICATION:  530
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Resnick, David S.
;       REGISTRATION NUMBER:  34,235
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (617)523-3400
;       TELEFAX:  (617)523-6440
;   INFORMATION FOR SEQ ID NO:  12:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  7 amino acids
;       TYPE:  amino acid
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  peptide
US-09-070-756-12


  Query Match             50.0%;  Score 22;  DB 2;  Length 7;
  Best Local Similarity   80.0%;  Pred. No. 1.7e+05;
  Matches    4;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       2 GRSFT 6
            || ||
Db       1 GREFT 5


RESULT   7
US-08-406-809-1
; Sequence 1, Application US/08406809
; Patent No. 5580788
;  GENERAL INFORMATION:
;     APPLICANT:  KINIKA, Yasunori
;     APPLICANT:  AIBA, Sachiko
;     TITLE OF INVENTION:  IMMUNOGLOBULIN BINDING ARTIFICIAL
;     TITLE OF INVENTION:  PROTEIN
;     NUMBER OF SEQUENCES:  11
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Browdy and Neimark
;       STREET:  419 Seventh St., NW, Suite 300
;       CITY:  Washington
;       STATE:  DC
;       COUNTRY:  USA
;       ZIP:  20004
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
```

```
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/09/499,909
;       FILING DATE:
;       CLASSIFICATION:  530
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/xxx,xxx
;       FILING DATE:  24-MAR-1993
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Tarcza, John E.
;       REGISTRATION NUMBER:  33,638
;       REFERENCE/DOCKET NUMBER:  Kilita 1
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  202 628 5197
;       TELEFAX:  202 737 3528
;       TELEX:  248633
;   INFORMATION FOR SEQ ID NO:  1:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  8 amino acids
;       TYPE:  amino acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  peptide
US-09-499-909-1


  Query Match             47.7%;  Score 21;  DB 1;  Length 8;
  Best Local Similarity   80.0%;  Pred. No. 1.7e+05;
  Matches    4;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 FGRSF 5
           | |||
Db       2 FARSF 6


RESULT   8
US-08-350-260A-469
; Sequence 469, Application US/08350260A
; Patent No. 5662255
;  GENERAL INFORMATION:
;    APPLICANT:  Winter, Gregory Paul
;    APPLICANT:  Griffiths, Andrew David
;    APPLICANT:  Williams, Samuel Cameron
;    APPLICANT:  Waterhouse, Peter
;    APPLICANT:  Nissim, Ahuva
;    APPLICANT:  Johnson, Kevin S.
;    APPLICANT:  Smith, Andrew John Hammond
;    TITLE OF INVENTION:  Methods for producing members of specific
;    TITLE OF INVENTION:  binding pairs
;    NUMBER OF SEQUENCES:  602
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  David W. Clough
;      STREET:  Marshall, O'Toole, Gerstein, Murray & Bicun
;      STREET:  6300 Sears Tower, 233 South Wacker Drive
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  USA
;      ZIP:  60606-6402
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25 (EPO)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/350,260A
;      FILING DATE:  05-DEC-1994
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  GB 9110547.4
;      FILING DATE:  15-MAY-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  GB 9206318.9
;      FILING DATE:  24-MAR-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/GB91/01134
;      FILING DATE:  10-JUL-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/GB92/00883
;      FILING DATE:  15-MAY-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/GB93/00605
;      FILING DATE:  24-MAR-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/150,002
;      FILING DATE:  21-MAR-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/307,619
;      FILING DATE:  16-SEP-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Clough, David W
;      REGISTRATION NUMBER:  36,107
;      REFERENCE/DOCKET NUMBER:  20111/32272
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312-474-6300
;  INFORMATION FOR SEQ ID NO:  469:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  8 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-350-260A-469


  Query Match             47.7%;  Score 21;  DB 2;  Length 8;
  Best Local Similarity   80.0%;  Pred. No. 1.7e+05;
  Matches    4;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 FGRSF 5
           | |||
Db       2 FARSF 6


RESULT   9
US-08-215-805A-54
; Sequence 54, Application US/08215805A
; Patent No. 5559008
;  GENERAL INFORMATION:
;    APPLICANT:  Chang, Yung-Fu
;    TITLE OF INVENTION:  LEUKOTOXIN GENE FROM PASTEURELLA
;    TITLE OF INVENTION:  SUIS
;    NUMBER OF SEQUENCES:  84
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Nixon, Hargrave, Devans & Doyle
;      STREET:  Clinton Square, P.O. Box 1051
;      CITY:  Rochester
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  14603
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/215,805A
;      FILING DATE:  22-MAR-1994
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Timian, Susan J.
;      REGISTRATION NUMBER:  34,103
;      REFERENCE/DOCKET NUMBER:  1965/61 (D 1329A)
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (716) 263 1636
;      TELEFAX:  (716) 263-1600
;  INFORMATION FOR SEQ ID NO:  54:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  9 amino acids
```

```
;       TYPE:  amino acid
;       STRANDEDNESS:
;       TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;     ORIGINAL SOURCE:
;       ORGANISM:  Actinobacillus pleuropneumoniae
US-08-213-897A-54


  Query Match             17.7%;  Score 27;  DB 1;  Length 7;
  Best Local Similarity   66.7%;  Pred. No. 1.7e+05;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    1;  Gaps    0;

Qy       1 FGRSFT 6
           ||  | |
Db       2 FGSKFT 7


RESULT  10
US-08-136-743B-54
; Sequence 54, Application US/08136743B
; Patent No. 5459063
;  GENERAL INFORMATION:
;    APPLICANT:  Barry S. Cooperman, Harvey Rubin,
;    APPLICANT:  Jerome Salom, and Alison L. Fisher
;    TITLE OF INVENTION:  "Plasmodium falciparum Ribonu
;    TITLE OF INVENTION:  cleotide Reductase, DNA Sequences Therefor and Peptide Inhibit
;    TITLE OF INVENTION:  Thereof"
;    NUMBER OF SEQUENCES:  67
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  The University of Pennsylvania
;      STREET:  Suite 320
;      STREET:  3700 Market Street
;      CITY:  Philadelphia
;      STATE:  Pennsylvania
;      COUNTRY:  U.S.A.
;      ZIP:  19104-3246
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.50 inch, 720 Kb
;      COMPUTER:  IBM PS/2
;      OPERATING SYSTEM:  MS-DOS
;      SOFTWARE:  WordPerfect 5.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/136,743B
;      FILING DATE:  12/31/93
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Monaco, Daniel A.
;      REGISTRATION NUMBER:  30,480
;      REFERENCE/DOCKET NUMBER:  3957-10
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (215) 568-8383
;      TELEFAX:  (215) 568-5549
;      TELEX:  845-0630
;  INFORMATION FOR SEQ ID NO:  54:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  7 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
US-08-136-743B-54


  Query Match             45.5%;  Score 20;  DB 1;  Length 7;
  Best Local Similarity   80.0%;  Pred. No. 1.7e-05;
  Matches    4;  Conservative    1;  Mismatches    0;  Indels    1;  Gaps    1;

Qy       5 FTLAS 9
           |:|:|
Db       1 FTIAA 5


RESULT  11
US-08-213-897A-5
; Sequence 5, Application US/08213897A
; Patent No. 5618790
;  GENERAL INFORMATION:
;    APPLICANT:
;    TITLE OF INVENTION:  Protease Mediated Iron Delivery System
;    NUMBER OF SEQUENCES:  18
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30 (EPO)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/213,897A
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/593,867
;      FILING DATE:  05-OCT-1990
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/833,283
;      FILING DATE:  10-FEB-1992
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  7 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-213-897A-5


  Query Match             45.5%;  Score 20;  DB 1;  Length 7;
  Best Local Similarity   57.1%;  Pred. No. 1.7e-05;
  Matches    4;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 FGRSFTL 7
           ||  | |
Db       1 FGHFFVL 7


RESULT  12
US-09-040-216-26
; Sequence 26, Application US/09040216
; Patent No. 6030942
; GENERAL INFORMATION:
;  APPLICANT: COOPERMAN, ET AL., BARRY
;  TITLE OF INVENTION: SMALL MOLECULES USEFUL FOR INHIBITING THE ACTIVITY OF
;  TITLE OF INVENTION: RIBONUCLEOTIDE REDUCTASE
;  FILE REFERENCE: 9596-630I
;  CURRENT APPLICATION NUMBER: US/09/040,216
;  CURRENT FILING DATE: 1998-03-17
;  EARLIER APPLICATION NUMBER: 08/919,748
;  EARLIER FILING DATE: 1997-08-28
;  EARLIER APPLICATION NUMBER: 60/025,149
;  EARLIER FILING DATE: 1996-08-30
;  NUMBER OF SEQ ID NOS: 61
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 26
;   LENGTH: 7
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence
;   OTHER INFORMATION: Ribonucleotide reductase inhibitor peptide
;   FEATURE:
;   OTHER INFORMATION: residue 1:  MET_RES:  ACTIVIAL N
US-09-040-216-26


  Query Match             45.5%;  Score 20;  DB 4;  Length 7;
  Best Local Similarity   80.0%;  Pred. No. 1.7e-05;
  Matches    4;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       5 FTLAS 9
```

```
                                                                        ;
Db       1 EFIAA 5


RESULT  14
US-08-846-333-45
; Sequence 45, Application US/08846333
; Patent No. 5777954
;  GENERAL INFORMATION:
;    APPLICANT:  Bernstein, David
;    APPLICANT:  Pulkkili, Timothy
;    TITLE OF INVENTION:  Methods for modifying DNA and I
;    TITLE OF INVENTION:  detecting effects of such modifications, & interaction of
;    TITLE OF INVENTION:  encoded modified polypeptides with target substrates.
;    NUMBER OF SEQUENCES:  86
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Richard F. Trecartin
;      STREET:  4 Embarcadero Center, Suite 3400
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94111
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/846,333
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/430,501
;      FILING DATE:
;      APPLICATION NUMBER:  US 07/602,158
;      FILING DATE:  22 OCT 1990
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Trecartin, Richard F.
;      REGISTRATION NUMBER:  31,801
;      REFERENCE/DOCKET NUMBER:  A-54369/RFT
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 781-1989
;      TELEFAX:  (415) 398-3249
;  INFORMATION FOR SEQ ID NO:  45:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  9 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-846-333-45


  Query Match             45.5%;  Score 20;  DB 5;  Length 9;
  Best Local Similarity   80.0%;  Pred. No. 1.7e+05;
  Matches    4;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       4 SFILA 8
           :||||
Db       1 EFIAA 5


RESULT  15
US-08-846-333-45
; Sequence 45, Application US/08846333
; Patent No. 5777954
;  GENERAL INFORMATION:
;    APPLICANT:  Bernstein, David
;    APPLICANT:  Pulkkili, Timothy
;    TITLE OF INVENTION:  Methods for modifying DNA and I
;    TITLE OF INVENTION:  detecting effects of such modifications, & interaction of
;    TITLE OF INVENTION:  encoded modified polypeptides with target substrates.
;    NUMBER OF SEQUENCES:  86
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Richard F. Trecartin
;      STREET:  4 Embarcadero Center, Suite 3400
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94111
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/846,333
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Trecartin, Richard F.
;      REGISTRATION NUMBER:  31,801
;      REFERENCE/DOCKET NUMBER:  A-54369/RFT
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 781-1989
;      TELEFAX:  (415) 398-3249
;  INFORMATION FOR SEQ ID NO:  45:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  9 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
PCT-US91-07506-45


  Query Match             45.5%;  Score 20;  DB 5;  Length 9;
  Best Local Similarity   80.0%;  Pred. No. 1.7e+05;
  Matches    4;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       4 SFILA 8
           :||||
Db       5 IFILA 9


RESULT  15
US-08-456-840-32
; Sequence 32, Application US/08456840
; Patent No. 5597908
;  GENERAL INFORMATION:
;    APPLICANT:  Taddei-Peters, W. C.
;    APPLICANT:  Butler, Sandra M.
;    TITLE OF INVENTION:  Immunoreactive Peptides of Apo(a)
;    NUMBER OF SEQUENCES:  48
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Akzo No. 5597908o1
;      STREET:  1330 Piccard Drive
;      CITY:  Rockville
;      STATE:  Maryland
;      COUNTRY:  US
;      ZIP:  20850
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC DOS/MS DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/475,340
;      FILING DATE:  01-JUN 1995
;      CLASSIFICATION:  536
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/266,407
;      FILING DATE:  27 JUN 1994
;      APPLICATION NUMBER:  US 08/172,461
;      FILING DATE:  21-DEC 1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Gormley, Mary E.
```

; REGISTRANT NUMBER: 84,409
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 301-258-5200
; TELEFAX: 301-977-0847
; INFORMATION FOR SEQ ID NO: 32:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-456-840-32

Query Match 43.2%; Score 19; DB 1; Length 5;
Best Local Similarity 66.8%; Prod. No. 1.2e+05;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PDSE 4
:::
Db 1 PDKHF 5

Search completed: August 26, 2002, 06:23:20
Job time: 761 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

SW protein protein search, using sw model

Run on: August 26, 2002, 06:24:48 ; Search time 7.9 Seconds
(without alignments)
11.702 Million cell updates/sec

Title: US-09-744-282-2
Perfect score: 43
Sequence: 1 FIASSEIG 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 788

Minimum DB seq length: 0
Maximum DB seq length: 9

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_71:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	28	65.1	7	2	PT0702	T-cell receptor be
2	28	65.1	9	2	PB0935	T-cell receptor be
3	25	58.1	6	2	PT0619	T-cell receptor be
4	23	48.8	6	2	PT0630	T-cell receptor be
5	21	48.8	6	2	PT0668	T-cell receptor be
6	21	48.8	7	2	PT0667	T-cell receptor be
7	21	48.8	7	2	PT0719	T-cell receptor be
8	21	48.8	9	2	PB0918	T-cell receptor be
9	20	46.5	6	2	PT0550	T-cell receptor be
10	20	46.5	7	2	PT0524	T-cell receptor be
11	20	46.5	7	2	PT0683	T-cell receptor be
12	20	46.5	8	2	PT0627	T-cell receptor be
13	20	46.5	8	2	PT0509	T-cell receptor be
14	19	44.2	5	2	PT0561	T-cell receptor be
15	18	41.9	5	2	PT0610	T-cell receptor be
16	18	41.9	6	2	PT0516	T-cell receptor be
17	18	41.9	6	2	PT0637	T-cell receptor be
18	18	41.9	6	2	PT0725	T-cell receptor be
19	18	41.9	7	2	PT0523	T-cell receptor be
20	18	41.9	7	2	PT0642	T-cell receptor be
21	18	41.9	9	2	PB0902	T-cell receptor be
22	17.5	40.7	8	2	A32523	peptidyl dipeptida
23	17.5	40.7	8	2	PT0595	T-cell receptor be
24	17	39.5	5	2	PT0535	T-cell receptor be
25	17	39.5	5	2	PT0684	T-cell receptor be
26	17	39.5	5	2	PT0565	T-cell receptor be
27	17	39.5	6	2	PT0518	T-cell receptor be
28	17	39.5	6	2	PT0648	T-cell receptor be
29	17	39.5	7	2	A25269	beta-glucosidase (AM3
30	17	39.5	7	2	PT0529	T-cell receptor be
31	17	39.5	9	2	PJ0676	T-cell receptor be
32	17	39.5	9	2	PB0917	T-cell receptor be
33	17	39.5	9	2	PB0921	T-cell receptor be
34	16	37.2	5	2	PT0686	T-cell receptor be
35	16	37.2	6	2	PT0687	T-cell receptor be
36	16	37.2	6	2	PT0662	T-cell receptor be
37	16	37.2	7	2	PT0696	T-cell receptor be
38	16	37.2	7	2	PT0655	T-cell receptor be
39	16	37.2	7	2	PT0544	T-cell receptor be
40	16	37.2	8	2	A25836	L-serine dehydrata
41	16	37.2	8	2	I48934	apolipoprotein A-I
42	16	37.2	8	2	PT0540	T-cell receptor be
43	16	37.2	8	2	PT0527	T-cell receptor be
44	16	37.2	8	2	PT0639	T-cell receptor be
45	16	37.2	8	2	PT0631	T-cell receptor be

ALIGNMENTS

RESULT 1
PT0702
T-cell receptor beta chain V-D-J region (PCT 2AA) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 01-May-1997
C;Accession: PT0702
R;Feeney, A.J.
J. Exp. Med. 174, 115-124, 1991
A;Title: Junctional sequences of total T cell receptor beta chains have few N regions
A;Reference number: PT0509; MUID:91277601
A;Accession: PT0702
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-7 <FEE>
A;Experimental source: newborn thymus; strain BALB/c
C;Keywords: T-cell receptor

Query Match 65.1%; Score 28; DB 2; Length 7;
Best Local Similarity 100.0%; Pred. No. 2.2e-05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 ASSEIG 9
||||||
Db 1 ASSEIG 6

RESULT 2
PB0935
T-cell receptor beta chain V-D-J region (clone 6) - rat (fragment)
C;Species: Rattus norvegicus (Norway rat)
C;Date: 09-Oct-1992 #sequence_revision 09-Oct-1992 #text_change 01-May-1997
C;Accession: PB0935
R;Gold, D.P.; Ottner, H.; Sun, D.; Wiley, S.; Vandenbark, A.A.; Wilson, L.B.
J. Exp. Med. 174, 1467-1476, 1991
A;Title: Analysis of T cell receptor beta chains in lewis rats with experimental allergic
A;Reference number: PB0891; MUID:92078857
A;Accession: PB0935
A;Molecule type: mRNA
A;Residues: 1-9 <GOL>
A;Experimental source: complete Freund's adjuvant immunized lymph node
C;Keywords: T-cell receptor

Query Match 65.1%; Score 28; DB 2; Length 9;
Best Local Similarity 100.0%; Pred. No. 2.2e-05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 ASSEIG 9
||||||
Db 2 ASSEIG 7

RESULT 5
PT0668
T-cell receptor beta chain V-D-J region (121 aa) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
C;Accession: PT0668
R;Feeney, A.J.
J. Exp. Med. 174, 115-124, 1991
A;Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.
A;Reference number: PT0509; MUID:91277601
A;Accession: PT0668
A;Status: translation not shown
A;Molecule type: mRNA
A;Residues: 1-6 <FEE>
A;Experimental source: day 4 postnatal thymus, strain BALB/c
C;Keywords: T-cell receptor

Query Match 48.8%; Score 21; DB 2; Length 6;
Best Local Similarity 83.3%; Pred. No. 2.8e-05;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 4 ASSEIG 9
:|||:|
Db 1 ASSGIG 6

RESULT 6
PT0667
T-cell receptor beta chain V-D-J region (121 aa) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
C;Accession: PT0667
R;Feeney, A.J.
J. Exp. Med. 174, 115-124, 1991
A;Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.
A;Reference number: PT0509; MUID:91277601
A;Accession: PT0667
A;Status: translation not shown
A;Molecule type: mRNA
A;Residues: 1-7 <FEE>
A;Experimental source: day 4 postnatal thymus, strain BALB/c
C;Keywords: T-cell receptor

Query Match 48.8%; Score 21; DB 2; Length 7;
Best Local Similarity 83.3%; Pred. No. 2.8e-05;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 4 ASSEIG 9
:|||:|
Db 1 ASSEGG 6

RESULT 7
PT0719
T-cell receptor beta chain V-D-J region (140-2F) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
C;Accession: PT0719; PT0638
R;Feeney, A.J.
J. Exp. Med. 174, 115-124, 1991
A;Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.
A;Reference number: PT0509; MUID:91277601
A;Accession: PT0719
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-7 <FEE>
A;Experimental source: newborn thymus, strain BALB/c (clone 140-2F)
A;Accession: PT0638
A;Status: translation not shown
A;Molecule type: mRNA
A;Residues: 1-7 <FEZ>
A;Experimental source: newborn thymus, strain BALB/c (clone 111-1N)
C;Keywords: T-cell receptor

Query Match 48.8%; Score 21; DB 2; Length 7;
Best Local Similarity 83.3%; Prod. No. 2.8e-05;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 4 ASSEIG 9
Db 1 ASSIIG 6

RESULT 8
PH0918
T cell receptor beta chain V-D-J region (isolate 4) rat (fragment)
C;Species: Rattus norvegicus (Norway rat)
C;Date: 30-Oct-1991 #sequence_revision 09-Oct-1992 #text_change 5-May-1997
C;Accession: PH0918
R;Gold, D.P.; Attner, H.; Sun, D.; Wiley, S.; Vandenbark, A.A.; Wilson, D.B.
J. Exp. Med. 174, 1467-1476, 1991
A;Title: Analysis of T cell receptor beta chains in Lewis rats with experimental allergi
A;Reference number: PH0891; MUID:92078857
A;Accession: PH0918
A;Molecule type: mRNA
A;Residues: 1-9 <GOL>
A;Experimental source: concanavalin A activated lymphoblast
C;Keywords: T-cell receptor

Query Match 48.8%; Score 21; DB 2; Length 9;
Best Local Similarity 83.3%; Prod. No. 2.8e-05;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 4 ASSEIG 9
Db 2 ASSGIG 7

RESULT 9
PI0550
T cell receptor beta chain V-D-J region (126 10F) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 5-May-1997
C;Accession: PI0550
R;Feeney, A.J.
J. Exp. Med. 174, 115-124, 1991
A;Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.
A;Reference number: PI0509; MUID:91277601
A;Accession: PI0550
A;Status: translation not shown
A;Molecule type: mRNA
A;Residues: 1-6 <FEE>
A;Experimental source: day 18 fetal thymus, strain BALB/c
C;Keywords: T-cell receptor

Query Match 46.5%; Score 20; DB 2; Length 6;
Best Local Similarity 66.7%; Prod. No. 2.8e-05;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 4 ASSEIG 9
Db 1 ASSIAG 6

RESULT 10
PI0524
T cell receptor beta chain V-D-J region (100-4AD) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
C;Accession: PI0524
R;Feeney, A.J.
J. Exp. Med. 174, 115-124, 1991
A;Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.
A;Reference number: PI0509; MUID:91277601
A;Accession: PI0524
A;Status: translation not shown
A;Molecule type: mRNA
A;Residues: 1-7 <FEE>
A;Experimental source: adult thymus, strain BALB/c
C;Keywords: T-cell receptor

Query Match 46.5%; Score 20; DB 2; Length 7;
Best Local Similarity 66.7%; Prod. No. 2.8e-05;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 4 ASSEIG 9
Db 1 ASSDAG 6

RESULT 11
PT0683
T cell receptor beta chain V-D-J region (240 1N) mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 5-May-1997
C;Accession: PT0683
R;Feeney, A.J.
J. Exp. Med. 174, 115-124, 1991
A;Title: Junctional sequences of fetal T cell receptor beta chains have few N regions
A;Reference number: PT0509; MUID:91277601
A;Accession: PT0683
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-7 <FEE>
A;Experimental source: day 18 fetal thymus, strain BALB/c
C;Keywords: T-cell receptor

Query Match 46.5%; Score 20; DB 2; Length 7;
Best Local Similarity 66.7%; Prod. No. 2.8e-05;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 4 ASSEIG 9
Db 1 ASSDAG 6

RESULT 12
PT0627
T cell receptor beta chain V-D-J region (100-2B) mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 5-May-1997
C;Accession: PT0627
R;Feeney, A.J.
J. Exp. Med. 174, 115-124, 1991
A;Title: Junctional sequences of fetal T cell receptor beta chains have few N regions
A;Reference number: PT0509; MUID:91277601
A;Accession: PT0627
A;Status: translation not shown
A;Molecule type: mRNA
A;Residues: 1-8 <FEE>
A;Experimental source: newborn thymus, strain BALB/c
C;Keywords: T-cell receptor

Query Match 46.5%; Score 20; DB 2; Length 8;
Best Local Similarity 66.7%; Prod. No. 2.8e-05;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 4 ASSEIG 9
Db 1 ASSDAG 6

RESULT 13
PT0509
T cell receptor beta chain V-D-J region (100-4N) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
C;Accession: PT0509
R;Feeney, A.J.

J. Exp. Med. 174, 115-124, 1991
A;Title: Functional sequences of fetal T cell receptor beta chains have few N regions.
A;Reference number: P10509; MUID:91277601
A;Accession: A;
A;Status: translation not shown
A;Molecule type: mRNA
A;Residues: 1-8<FEE>
A;Experimental source: adult thymus, strain BALB/c
C;Keywords: T-cell receptor

Query Match 46.5%; Score 20; DB 2; Length 8;
Best Local Similarity 66.7%; Pred. No. 2.4e+05;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ASSEE 5
:::
Db 1 ASSAG 5

RESULT 14
P10561
T-cell receptor beta-chain V-D-J region (126-11) mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 01-May-1997
C;Access: A; P10561
R;Freedy, A.J.
J. Exp. Med. 174, 115-124, 1991
A;Title: Functional sequences of fetal T cell receptor beta chains have few N regions.
A;Reference number: P10509; MUID:91277601
A;Accession: A; P10561
A;Status: translation not shown
A;Molecule type: mRNA
A;Residues: 1-5<FEE>
A;Experimental source: day 18 fetal thymus, strain BALB/c
C;Keywords: T-cell receptor

Query Match 44.2%; Score 19; DB 2; Length 5;
Best Local Similarity 80.0%; Pred. No. 2.8e+05;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ASSEE 5
:::
Db 1 ASSIC 5

RESULT 15
P10560
T-cell receptor beta-chain V-D-J region (100.2E) mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 01-May-1997
C;Access: A; P10560
R;Freedy, A.J.
J. Exp. Med. 174, 115-124, 1991
A;Title: Functional sequences of fetal T cell receptor beta chains have few N regions.
A;Reference number: P10509; MUID:91277601
A;Accession: A; P10560
A;Status: translation not shown
A;Molecule type: mRNA
A;Residues: 1-5<FEE>
A;Experimental source: newborn thymus, strain BALB/c
C;Keywords: T-cell receptor

Query Match 41.9%; Score 18; DB 2; Length 5;
Best Local Similarity 80.0%; Pred. No. 2.8e+05;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ASSEE 5
:::
Db 1 ASSIC 5

Search completed: August 26, 2002, 06:24:48
Job time: 849 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 26, 2002, 06:40:42 ; Search time 44.94 Seconds
(without alignments)
7.754 Million cell updates/sec

Title: US-09-744-282-2
Perfect score: 43
Sequence: 1 FILASSEIG 9

Scoring table: BLOSUM62
Gapop 13.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 231

Minimum DB seq length: 0
Maximum DB seq length: 9

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	17	39.5	7	1	CIA_ENTFA	P11972 enterococcus
2	14	32.6	8	1	UF06_MOUSE	P38644 mus musculus
3	14	32.6	9	1	THYE_PIG	P01595 sus scrofa
4	13	30.2	9	1	DI_NEPNE	P24816 nephrops no
5	13	30.2	9	1	DSIP_RABIT	P01156 oryctolagus
6	12	27.9	8	1	RS1_ERWCH	P27985 erwinia chr
7	11	25.6	5	1	RE41_LITRU	P80..72 litoria rub
8	11	25.6	5	1	RE42_LITRU	P80.73 litoria rub
9	11	25.6	7	1	UHI1_RAT	P06576 rattus norv
10	11	25.6	8	1	AKBG_GRYBI	P14796 gryllus bim
11	11	25.6	8	1	AKH_LIBAU	P29418 libellula a
12	11	25.6	8	1	AKH_TABAT	P24995 tabanus atr
13	11	25.6	8	1	CAD1_ENTFA	P13208 enterococcus
14	11	25.6	8	1	BIF2_PERAM	P34539 periplaneta
15	11	25.6	8	1	LMT2_LOCMI	P22396 locusta mig
16	11	25.6	8	1	LPK_LEUMA	P15969 leucophaea
17	11	25.6	8	1	ORMY_ORCLI	P08465 orconectes
18	11	25.6	9	1	ALIT_CARMA	P81833 carcinus ma
19	11	25.6	9	1	CCAP_CARMA	P30956 carcinus ma
20	11	25.6	9	1	LITR_PHYRO	P08949 phyllomedus
21	11	25.6	9	1	OXYI_EISFO	P42908 eisenia foe
22	11	25.6	9	1	ULAD_HUMAN	P31929 homo sapien
23	10	23.3	6	1	LOK1_LOCMI	P41341 locusta mig
24	10	23.3	7	1	GFRP_MOUSE	P99026 mus musculu
25	10	23.3	8	1	FUSS_FUSSO	P81010 fusarium so
26	10	23.3	8	1	LCK7_LEUMA	P15989 leucophaea
27	10	23.3	9	1	FAR6_CALVO	P41961 calliphora
28	10	23.3	9	1	FLA2_TREHY	P80259 treponema h
29	10	23.3	9	1	LPCA_STAAU	P09884 staphylococ
30	10	23.3	9	1	ULAH_HUMAN	P31934 homo sapien
31	10	23.3	9	1	XYLA_STRSQ	P15349 streptomyce
32	9	20.9	4	1	EOSI_HUMAN	P82..52 homo sapien
33	9	20.9	5	1	BIOB_CITFR	P12997 citrobacter
34	9	20.9	7	1	CCF1_ENTFA	P..0104 enterococcus
35	9	20.9	7	1	IGAC_DAVDE	P..6264 dactylium d
36	9	20.9	7	1	LANC_CAKUT	P..9960 sarcobacter
37	9	20.9	8	1	GLUR_HUMAN	P..2729 homo sapien
38	9	20.9	9	1	TAL1_PICJA	P..7440 pichia jadi
39	9	20.9	9	1	TAL4_PICJA	P..7441 pichia jadi
40	9	20.9	9	1	UNI9_CLOPA	P..1859 clostridium
41	9	20.9	9	1	YBER_AZOVI	P..5826 azotobacter
42	8	18.6	5	1	AL14_CARMA	P..0617 carcinus ma
43	8	18.6	5	1	PSK_DAUCA	P..8261 daucus caro
44	8	18.6	7	1	AL12_CARMA	P..1805 carcinus ma
45	8	18.6	7	1	AL14_CARMA	P..1806 carcinus ma

ALIGNMENTS

RESULT 1
CIA_ENTFA
ID CIA_ENTFA STANDARD; PRT; 7 AA.
AC P11972;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 01-FEB-1991 (Rel. 17, Last annotation update)
DE Sex pheromone cAM373 (clumping-inducing agent) (CIA).
OS Enterococcus faecalis (Streptococcus faecalis).
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Enterococcaceae;
OC Enterococcus.
OX NCBI_TaxID=1351;
RN [1]
RP SEQUENCE.
RX MEDLINE=87005252; PubMed=3093276;
RA Mori M., Tanaka H., Sakakami Y., Isomi A., Fujii M., Kitada C.,
RA White B.A., An F.Y., Clewell D.B., Suzuki A.;
RT "Isolation and structure of the Streptococcus faecalis sex pheromone,
RT cAM373.";
RL FEBS Lett. 206:69-72(1986).
CC -!- FUNCTION: cAM373 INDUCES MATING RESPONSE IN P-R NEAR CELLS
CC HARBORING pAM373.
CC -!- MISCELLANEOUS: THE N-TERMINUS IS POSSIBLY RESPONSIBLE FOR
CC SPECIFICITY OF PHEROMONES TO PLASMIDS.
CC -!- SIMILARITY: C-TERMINAL 5 IBG PHEROMONES ARE ALL ACID.
DR PIR; A25269; A25269.
KW Pheromone.
SQ SEQUENCE 7 AA; 734 MW; 75BD87205900505E0 CRC64;

Query Match 39.5%; Score 17; DB 1; Length 7;
Best Local Similarity 80.0%; Pred. No. 1e-05;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FILAS 5
| |||
Db 3 FILAS 7

RESULT 2
UF06_MOUSE
ID UF06_MOUSE STANDARD; PRT; 8 AA.
AC P38644;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 01-FEB-1995 (Rel. 31, Last annotation update)
DE Unknown protein from 2D page of fibroblasts (P5) (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Cephalostomi;
OC Mammalia; Eutheria; Rodentia; Sciuromorpha; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE.
RC TISSUE=Fibroblast;
RX MEDLINE=95009907; PubMed=7525104;
RA Merrick B.A., Patterson K.M., Wichter J.A., Bo..., Selkirk J.K.,

```
RL   "Separation and sequencing of familiar and novel marine proteins
RL   using preparative two-dimensional gel electrophoresis.";
RL   Electrophoresis 15:735-746(1994).
CC   -!- MISCELLANEOUS: ON THE 2D GEL THE DETERMINED PI OF THIS UNKNOWN
CC       FOR DESN IS: 5.2, ITS MW IS: 50 KDa.
DR   N_NEMER; 8; 8.
SQ   SEQUENCE   9 AA;  817 MW;  A45B8B78B2B6B0B0B1 CRC64;


  Query Match             42.9%;  Score 14;  DB 1;  Length 9;
  Best Local Similarity   75.0%;  Pred. No. 1e+05;
  Matches    3;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       5 SEDG 9
           | ||
Db       2 SEAG 5


RESULT   4
DSIP_HUMAN
ID   DSIP_HUMAN      STANDARD;      PRT;     9 AA.
AC   P01157;
DT   21-JUL-1986 (Rel. 01, Created)
DT   21-JUL-1986 (Rel. 01, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Delta sleep-inducing peptide (DSIP).
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Primates; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=77185324; PubMed=862769;
RA   Schoenenberger G.A., Maier P.F., Tobler H.J., Wilson K., Mueller M.;
RT   "The delta sleep inducing peptide (DSIP). Comparative properties of
RT   the original and synthetic nonapeptide.";
RL   Experientia 33:548-552(1977).
RN   [2]
RP   SEQUENCE, AND SYNTHESIS.
RX   MEDLINE=79054421; PubMed=568769;
RA   Schoenenberger G.A., Maier P.F., Tobler H.J., Wilson K., Mueller M.;
RT   "The delta EEG (sleep)-inducing peptide (DSIP). XI. Amino acid
RT   analysis, sequence, synthesis and activity of the nonapeptide.";
RL   Pflugers Arch. 376:119-129(1978).
RN   [3]
RP   REVIEW.
RX   MEDLINE=87175129; PubMed=3550726;
RA   Graf M.V., Kastin A.J.;
RT   "Delta-sleep-inducing peptide (DSIP): an update.";
RL   Peptides 7:1165-1187(1986).
CC   -!- FUNCTION: WHEN INFUSED INTO THE MESODIENCEPHALIC VENTRICLE OF
CC       RECIPIENT RABBITS INDUCES SPINDLE AND DELTA EEG ACTIVITY AND
CC       REDUCED MOTOR ACTIVITIES.
CC   -!- MISCELLANEOUS: THIS PEPTIDE WAS OBTAINED FROM DIALYSATES OF
CC       OCCIPITAL VENOUS SINUS BLOOD FROM RABBITS KEPT ASLEEP BY ELECTRIC
CC       STIMULATION OF THE THALAMUS.
CC   -!- DATABASE: NAME Protein Spotlight;
CC       NOTE=Issue 8 of March 2001;
CC       WWW="http://www.expasy.ch/spotlight/".
DR   PIR; A01422; QDRB.
SQ   SEQUENCE   9 AA;  849 MW;  DDD465BDDAA8787D CRC64;


  Query Match             30.2%;  Score 13;  DB 1;  Length 9;
  Best Local Similarity   75.0%;  Pred. No. 1e+05;
  Matches    3;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       4 ASSE 7
           | || |
Db       6 ASGE 9


RESULT   6
```

```
RS1_ERWCH
ID   RS1_ERWCH      STANDARD;      PRT;     8 AA.
AC   P47986;
DT   01-OCT-1994 (Rel. 30, Created)
DT   01-OCT-1994 (Rel. 30, Last sequence update)
DT   01-NOV-1995 (Rel. 32, Last annotation update)
DE   30S ribosomal protein S1 (Fragment).
GN   RPSA.
OS   Erwinia chrysanthemi.
OC   Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC   Pectobacterium.
OX   NCBI_TaxID=556;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=3937;
RA   Dreyfus A., Toussaint A., Faelen M.;
RL   Submitted (AUG 1993) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: BINDS MRNA; THUS FACILITATING RECOGNITION OF THE
CC       INITIATION REGION. IT IS NEEDED FOR TRANSLATION MRNA WITH A HIGH
CC       SHINE-DALGARNO (SD) PURINE-RICH SEQUENCE (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE S1P FAMILY OF RIBOSOMAL PROTEINS.
CC   ---------- --- ------ -- ---------- ---------- - --- ---------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------- --- ------ -- ---------- ---------- - --- ---------
DR   EMBL; X74750; CAA52769.1; -.
DR   PIR; S47141; S47141.
KW   Ribosomal protein; Repeat; RNA-binding.
FT   NON_TER       1       1
SQ   SEQUENCE   8 AA;  847 MW;  9E18743EC5B349CD CRC64;


  Query Match             27.9%;  Score 12;  DB 1;  Length 8;
  Best Local Similarity   42.9%;  Pred. No. 1e+05;
  Matches    3;  Conservative    0;  Mismatches    4;  Indels    1;  Gaps    0;

Qy       1 FILASSE 7
           |  |  :
Db       2 FKSAKDE 8


RESULT   7
RE41_LITRU
ID   RE41_LITRU     STANDARD;      PRT;     5 AA.
AC   P82072;
DT   01-MAR-2002 (Rel. 41, Created)
DT   01-MAR-2002 (Rel. 41, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
DE   Rubellidin 4.1.
OS   Litoria rubella (Desert tree frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Neobatrachia; Bufonoidea; Hylidae;
OC   Litoria.
OX   NCBI_TaxID=104895;
RN   [1]
RP   SEQUENCE, AND MASS SPECTROMETRY.
RC   TISSUE=Skin secretion;
RA   Steinborner S.I., Wabnitz P.A., Waugh R.J., Bowie J.H., Tyler M.J.,
RA   Tyler M.J., Wallace J.C.;
RT   "The structure of new peptides from the Australian red tree frog
RT   'Litoria rubella': the skin peptide profile as a probe for the study
RT   of evolutionary trends of amphibians.";
RL   Aust. J. Chem. 49:955-963(1996).
CC   -!- FUNCTION: SHOW NEITHER NEUROPEPTIDE ACTIVITY N R ANTIBIOTIC
CC       ACTIVITY.
CC   -!- TISSUE SPECIFICITY: SECRETED BY THE SKIN DORSAL GLANDS.
CC   -!- MASS SPECTROMETRY: MW=655; METHOD=FAB.
KW   Amphibian skin; Amidation.
FT   MOD_RES       5       5       AMIDATION.
SQ   SEQUENCE   5 AA;  656 MW;  71A9C9CB1C40C00C CRC64;


  Query Match             25.6%;  Score 11;  DB 1;  Length 5;
  Best Local Similarity   100.0%;  Pred. No. 1e+05;
  Matches    2;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 FT 2
             ||
Db       4 FT 5


RESULT   8
RE32_LITRU
ID   RE32_LITRU     STANDARD;      PRT;     5 AA.
AC   P82073;
DT   01-MAR-2002 (Rel. 41, Created)
DT   01-MAR-2002 (Rel. 41, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
DE   Rubellidin 3.2.
OS   Litoria rubella (Desert tree frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Neobatrachia; Bufonoidea; Hylidae;
OC   Litoria.
OX   NCBI_TaxID=104895;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Skin secretion;
RA   Wabnitz P.A., Bowie J.H., Tyler M.J., Wallace J.C.;
RT   "Peptides from the skin glands of the Australian bazzini tree frog
RT   Litori electrica. Comparison with the skin peptides from Litoria
RT   rubella.";
RL   Aust. J. Chem. 52:0-0(1999).
CC   -!- FUNCTION: SHOW NEITHER NEUROPEPTIDE ACTIVITY N R ANTIBIOTIC
CC       ACTIVITY.
CC   -!- TISSUE SPECIFICITY: SECRETED BY THE SKIN DORSAL GLANDS.
KW   Amphibian skin.
SQ   SEQUENCE   5 AA;  570 MW;  71A9C9CB2A0C0C1C CRC64;


  Query Match             25.6%;  Score 11;  DB 1;  Length 5;
  Best Local Similarity   100.0%;  Pred. No. 1e+05;
  Matches    2;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 FT 2
             ||
Db       4 FT 5


RESULT   9
UBIL_RAT
ID   UBIL_RAT       STANDARD;      PRT;     7 AA.
AC   P56576;
DT   15-DEC-1998 (Rel. 37, Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   15-DEC-1998 (Rel. 37, Last annotation update)
DE   Unknown protein from 2D-page of heart tissue (Spot PI1) (Fragment).
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE.
RC   STRAIN=WISTAR; TISSUE=Heart;
RA   Li X. P., Ploessner K. P., Scholer C., Boettger Zarresse V., Sallkow J.,
RA   Jungblut P.R.;
RL   Submitted (SEP 1998) to the SWISS-PROT data bank.
CC   -!- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI FOR UNKNOWN
CC       PROTEIN IS: 8.5, HIS MW IS: 42 KDa.
FT   UNSURE        2       2       OR A.
FT   NON_TER       7       7
SQ   SEQUENCE   7 AA;  775 MW;  6B6D4B4041E75A4A0 CRC64;
```

```
  Query Match             25.6%;  Score 11;  DB 1;  Length 8;
  Best Local Similarity   66.7%;  Pred. No. 1e+05;
  Matches    2;  Conservative    2;  Mismatches    2;  Indels    1;  Gaps    0;

Qy       3 ASGFL 7
           |:| :
Db       2 GAFPG 7


RESULT 10
AKH1_GRYBI
ID   AKH1_GRYBI      STANDARD;      PRT;     8 AA.
AC   P81066;
DT   01-JAN-1998 (Rel. 36, Created)
DT   01-FEB-1994 (Rel. 28, Last sequence update)
DT   01-FEB-1994 (Rel. 28, Last annotation update)
DE   Adipokinetic hormone G (AKH G) (RO II).
OS   Gryllus bimaculatus (Two spotted cricket), and
OS   Romalea microptera (Lubber grasshopper).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Orthopteroidea; Orthoptera; Ensifera; Gryllidae;
OC   Gryllinae; Gryllus.
OX   NCBI_TaxID=6999, 7007;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Corpora cardiaca;
RX   MEDLINE=86160654; PubMed=6226916;
RA   Gade G., Rinehart K.L. Jr.;
RT   "Primary sequence analysis by fast atom bombardment mass spectrometry
RT   of a peptide with adipokinetic activity from the corpora cardiaca of
RT   the cricket Gryllus bimaculatus.";
RL   Int. Journ. Biolgy. Res. Common. 14:968-974(1986).
RN   [2]
RP   SEQUENCE.
RC   TISSUE=Lepidoptera; TISSUE Corpora cardiaca;
RX   MEDLINE=88147023; PubMed=3226948;
RA   Gade G., Hilbich C., Beyreuther K., Rinehart K.L. Jr.;
RT   "Sequence analyses of two neuropeptides of the AKH/RPCH-family from
RT   the lubber grasshopper, Romalea microptera.";
RL   Peptides 9:681-688(1988).
CC   -!- FUNCTION: THIS HORMONE, RELEASED FROM CELLS IN THE CORPORA
CC       CARDIACA AFTER THE BEGINNING OF FLIGHT, CAUSES RELEASE OF
CC       DIGLYCERIDES FROM THE FAT BODY AND THEN STIMULATES THE FLIGHT
CC       MUSCLES TO USE THESE DIGLYCERIDES AS AN ENERGY SOURCE.
CC   -!- SIMILARITY: BELONGS TO THE AKH / HRTH / RPCH FAMILY.
DR   PIR; A28004; A28004.
DR   InterPro; IPR002047; AKH.
DR   PROSITE; PS00256; AKH; 1.
KW   Neuropeptide; Amidation; Flight.
FT   MOD_RES         1       1       PYRROLIDONE CARBOXYLIC ACID.
FT   MOD_RES         8       8       AMIDATION.
SQ   SEQUENCE   8 AA;  938 MW;  867861B5B9C45206 CRC64;


  Query Match             25.6%;  Score 11;  DB 1;  Length 8;
  Best Local Similarity   100.0%;  Pred. No. 1e+05;
  Matches    2;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       8 FT 9

Db       4 FT 5


RESULT 11
AKH_LIBAT
ID   AKH_LIBAT       STANDARD;      PRT;     8 AA.
AC   P23453;
DT   01-NOV-1991 (Rel. 22, Created)
DT   01-FEB-1994 (Rel. 28, Last sequence update)
DT   01-FEB-1994 (Rel. 28, Last annotation update)
DE   Adi[illegible] hormone (AKH).
```

```
OS   Libellula auripennis (Skimmer dragonfly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Palaeoptera; Odonata; Anisoptera; Libellulidae; Libellula.
OX   NCBI_TaxID=6966;
RN   [1]
RP   SEQUENCE, AND SYNTHESIS.
RC   TISSUE=Corpora cardiaca;
RX   MEDLINE=90359055; PubMed=2390214;
RA   Gade G.;
RT   "The putative ancestral peptide of the adipokinetic/red pigment
RT   concentrating hormone family isolated and sequenced from a
RT   dragonfly.";
RL   Biol. Chem. Hoppe Seyler 371:475-483(1990).
CC   -!- FUNCTION: THIS HORMONE, RELEASED FROM CELLS IN THE CORPORA
CC       CARDIACA AFTER THE BEGINNING OF FLIGHT, CAUSES RELEASE OF
CC       DIGLYCERIDES FROM THE FAT BODY AND THEN STIMULATES THE FLIGHT
CC       MUSCLES TO USE THESE DIGLYCERIDES AS AN ENERGY SOURCE.
CC   -!- SIMILARITY: BELONGS TO THE AKH / HRTH / RPCH FAMILY.
DR   PIR; S10596; S10596.
DR   InterPro; IPR002047; AKH.
DR   PROSITE; PS00256; AKH; 1.
KW   Neuropeptide; Amidation; Flight.
FT   MOD_RES         1       1       PYRROLIDONE CARBOXYLIC ACID.
FT   MOD_RES         8       8       AMIDATION.
SQ   SEQUENCE   8 AA;  978 MW;  8665A771A9C45206 CRC64;


  Query Match             25.6%;  Score 11;  DB 1;  Length 8;
  Best Local Similarity   100.0%;  Pred. No. 1e+05;
  Matches    2;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 FT 2
           ||
Db       4 FT 5


RESULT 12
AKH_TABAT
ID   AKH_TABAT       STANDARD;      PRT;     8 AA.
AC   P14595;
DT   01-JAN-1990 (Rel. 13, Created)
DT   01-FEB-1994 (Rel. 28, Last sequence update)
DT   01-FEB-1994 (Rel. 28, last annotation update)
DE   Adipokinetic hormone (AKH) (Dipteran corpora cardiaca factor 1)
DE   (DCC 1).
OS   Tabanus atratus (Horse fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Tabanomorpha;
OC   Tabanidae; Tabanus.
OX   NCBI_TaxID=7207;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Corpora cardiaca;
RX   MEDLINE=90046758; PubMed=2813485;
RA   Jaffe H., Raina A.K., Riley C.T., Fraser B.A., Nachman R.J.,
RA   Vogel V.W., Zhang Y. S., Bayes D.K.;
RT   "Primary structure of two neuropeptide hormones with adipokinetic and
RT   hypotrehalosemic activity isolated from the corpora cardiaca of horse
RT   flies (Diptera).";
RL   Proc. Natl. Acad. Sci. U.S.A. 86:8161-8164(1989).
CC   -!- FUNCTION: THIS HORMONE, RELEASED FROM CELLS IN THE CORPORA
CC       CARDIACA AFTER THE BEGINNING OF FLIGHT, CAUSES RELEASE OF
CC       DIGLYCERIDES FROM THE FAT BODY AND THEN STIMULATES THE FLIGHT
CC       MUSCLES TO USE THESE DIGLYCERIDES AS AN ENERGY SOURCE.
CC   -!- SIMILARITY: BELONGS TO THE AKH / HRTH / RPCH FAMILY.
DR   PIR; A33995; A33995.
DR   InterPro; IPR002047; AKH.
DR   PROSITE; PS00256; AKH; 1.
KW   Neuropeptide; Amidation; Flight.
FT   MOD_RES         1       1       PYRROLIDONE CARBOXYLIC ACID.
FT   MOD_RES         8       8       AMIDATION.
SQ   SEQUENCE   8 AA;  949 MW;  8678677[illegible]A9D1A736 CRC64;
```

```
  Query Match             25.6%;  Score 11;  DB 1;  Length 8;
  Best Local Similarity   100.0%;  Pred. No. 1e+05;
  Matches    2;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 FT 2
           ||
Db       4 FT 5


RESULT  13
CAD1_ENTFA
ID   CAD1_ENTFA     STANDARD;      PRT;     8 AA.
AC   P14268;
DT   01-JAN-1990 (Rel. 13, Created)
DT   01-JAN-1990 (Rel. 13, Last sequence update)
DT   01-FEB-1991 (Rel. 17, Last annotation update)
DE   Sex pheromone CAD1.
OS   Enterococcus faecalis (Streptococcus faecalis).
OC   Bacteria; Firmicutes; Bacillus/Clostridium group; Enterococcus group;
OC   Enterococcus.
OX   NCBI_TaxID=1351;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=85051889; PubMed=6437872;
RA   Mori M., Sakurai Y., Narita M., Isobai A., Fujii T., Kitada C.,
RA   Craig R.A., Clewell D.B., Suzuki A.;
RT   "Isolation and structure of the bacterial sex pheromone, cAM1, that
RT   induces plasmid transfer in Streptococcus faecalis.";
RL   FEBS Lett. 178:97-100(1984).
CC   -!- FUNCTION: CAD1 IS INVOLVED IN THE CONJUGATIVE TRANSFER OF THE
CC       HEMOLYSIN PLASMID PAD1.
KW   Pheromone.
SQ   SEQUENCE   8 AA;  819 MW;  047DD732C735B0C7 CRC64;


  Query Match             25.6%;  Score 11;  DB 1;  Length 8;
  Best Local Similarity   66.7%;  Pred. No. 1e+05;
  Matches    2;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 FTL 3
           |:|
Db       2 FSL 4


RESULT  14
HIF2_PERAM
ID   HIF2_PERAM     STANDARD;      PRT;     8 AA.
AC   P04844;
DT   15-AUG-1988 (Rel. 05, Created)
DT   01-FEB-1994 (Rel. 28, Last sequence update)
DT   01-FEB-1994 (Rel. 28, Last annotation update)
DE   Hypertrehalosaemic factor II (Neuropeptide M-II) (Periplanetin CC-2)
DE   (Pea-CAH-II) (Leu-CC-II) (Hypertrehalosaemic neuropeptide II).
OS   Periplaneta americana (American cockroach),
OS   Leptinotarsa decemlineata (Colorado potato beetle), and
OS   Blatta orientalis (oriental cockroach).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Orthopteroidea; Dictyoptera; Blattaria;
OC   Blattoidea; Blattidae; Periplaneta.
OX   NCBI_TaxID=6978, 7539, 6976;
RN   [1]
RP   SEQUENCE.
RC   SPECIES P.americana;
RX   MEDLINE 85046530; PubMed=6548628;
RA   Witten J.L., Schaffer M.H., Ohshio M., Orchi J.C., Bowman M.F.,
RA   Reichart K.L. JR.;
RT   "Structures of two cockroach neuropeptides assigned by fast atom
RT   bombardment mass spectrometry.";
RL   Biochem. Biophys. Res. Commun. 124:450-458(1984).
RN   [2]
RP   SEQUENCE.
RC   SPECIES P.americana
```

```
RX   MEDLINE 84298179; PubMed 6591205;
RA   Scarborough R.M., Jamieson G.C., Kalish P., Kramer S.J., McEnroe G.A.,
RA   Miller C.A., Schooley D.A.;
RT   "Isolation and primary structure of two peptides with
RT   cardioacceleratory and hyperglycemic activity from the corpora
RT   cardiaca of Periplaneta americana.";
RL   Proc. Natl. Acad. Sci. U.S.A. 81:5575-5579(1984).
RN   [3]
RP   SEQUENCE.
RC   SPECIES L.decemlineata; TISSUE Corpora cardiaca;
RX   MEDLINE 90160053; PubMed=2576128;
RA   Gaede G., Kellner R.;
RT   "The metabolic neuropeptides of the corpus cardiacum from the potato
RT   beetle and the American cockroach are identical.";
RL   Peptides 10:1287-1289(1989).
RN   [4]
RP   SEQUENCE.
RC   SPECIES B.orientalis; TISSUE Corpora cardiaca;
RX   MEDLINE 90253659; PubMed=2340112;
RA   Gaede G., Pinehart K.L. Jr.;
RT   "Primary structures of hypertrehalosaemic neuropeptides isolated from
RT   the corpora cardiaca of the cockroaches Leucophaea maderae,
RT   Gromphadorhina portentosa, Blattella germanica and Blatta orientalis
RT   and of the stick insect Extatosoma tiaratum assigned by tandem fast
RT   atom bombardment mass spectrometry.";
RL   Biol. Chem. Hoppe-Seyler 371:847-854(1990).
CC   -!- FUNCTION: HYPERTREHALOSAEMIC FACTORS ARE NEUROPEPTIDES THAT
CC       ELEVATE THE LEVEL OF TREHALOSE IN THE HEMOLYMPH (TREHALOSE IS
CC       THE MAJOR CARBOHYDRATE IN THE HEMOLYMPH OF INSECTS)
CC   -!- SIMILARITY: BELONGS TO THE ANA / CRZ / HCP FAMILY.
DR   PIR; A05170; A05170.
DR   PIR; S08996; S08996.
DR   PIR; B44960; B44960.
DR   PIR; B49823; B49823.
DR   InterPro; IPR002047; AKH.
DR   PROSITE; PS00256; AKH; 1.
KW   Neuropeptide; Amidation.
FT   MOD_RES       1       1       PYROGLUTAMIC ACID (PAR-XYL).
FT   MOD_RES       8       8       AMIDATION.
SQ   SEQUENCE   8 AA;  1006 MW;  B0745771A0D1A7A0 CRC64;


  Query Match             25.6%;  Score 11;  DB 1;  Length 8;
  Best Local Similarity   100.0%;  Pred. No. 1e+05;
  Matches    2;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 FT 2
           ||
Db       4 FT 5


RESULT  15
LMT2_LOCMI
ID   LMT2_LOCMI     STANDARD;      PRT;     8 AA.
AC   P22396;
DT   01-AUG-1991 (Rel. 19, Created)
DT   01-AUG-1991 (Rel. 19, Last sequence update)
DT   01-AUG-1991 (Rel. 19, Last annotation update)
DE   Locustamyotropin 2 (LOM-MT 2).
OS   Locusta migratoria (Migratory locust).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Orthopteroidea; Orthoptera; Caelifera;
OC   Acridomorpha; Acridoidea; Acrididae; Locusta.
OX   NCBI_TaxID=7004;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Corpora cardiaca;
RA   Schoofs L., Holman G.M., Hayes T.K., Nachman R.J., De Loof A.;
RT   "Isolation, identification and synthesis of locustamyotropin II, an
RT   additional neuropeptide of Locusta migratoria. Members of the
RT   cephalomyotropic peptide family.";
RL   Insect Biochem. 20:479-484(1990).
CC   -!- FUNCTION: MYOTROPIC VISCERAL MUSCLE - NERVE ACTIVITY
```

XX (MYOTROPIC ACTIVITY).
XX 1. ASSOCIATED WITH: MEMBERS OF THE PYROKININ FAMILY.
DE Leucokin-I, CRP-like; Pyrokinin.
DE Periplaneta americana PYROKININ-I.
KW Neuropeptide; Amidation; Pyrokinin.
FT MOD_RES 9 9 AMIDATION.
SQ SEQUENCE 9 AA; 944 MW; 2E441F7A1A9CAA8AB CRC64;

Query Match 25.0%; Score 11; DB 1; Length 9;
Best Local Similarity 100.0%; Pred. No. 1e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 8 FG 9

Db 3 FG 4

Search completed: August 26, 2002, 06:40:43

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 26, 2002, 06:39:44 ; Search time 135.92 Seconds
(without alignments)
11.455 Million cell updates/sec

Title: US-09-744-282-2
Perfect score: 43
Sequence: 1 FTLASSETG 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0
Maximum DB seq length: 9

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_19:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	21	48.8	9	15	Q85710	Q85710 rous sarcom
2	17	39.5	8	3	P82858	P82858 paecilia re
3	17	39.5	9	7	Q9MW43	Q9mw43 homo sapien
4	16	37.2	8	2	P72221	P72221 pseudomonas
5	16	37.2	8	11	Q60615	Q60615 mus musculu
6	16	37.2	9	2	Q935G0	Q935g0 salmonella
7	16	37.2	9	6	Q9XSL0	Q9xsl0 capra hircu
8	16	37.2	9	11	Q9QVH9	Q9qvh9 mus sp. sup
9	16	37.2	9	11	Q9QZA8	Q9qza8 mus musculu
10	15	34.9	7	2	P70804	P70804 azotobacter
11	15	34.9	8	4	Q9P285	Q9p285 homo sapien
12	15	34.9	8	5	Q94623	Q94623 manduca sex
13	15	34.9	9	12	Q66545	Q66545 human herpe
14	14	32.6	8	12	Q83977	Q83977 influenza a
15	14	32.6	9	5	Q9VV82	Q9vv82 drosophila
16	14	32.6	9	8	Q94V18	Q94v18 caratus oro
17	14	32.6	9	8	Q94VH4	Q94vh4 varanus gra
18	12	27.9	8	3	Q9HDS4	Q9hds4 aspergillus
19	12	27.9	8	8	Q9MSX1	Q9msx1 juriena fam
20	12	27.9	8	11	O35835	O35835 rattus sp.
21	12	27.9	9	2	Q9K4M6	Q9k4m6 staphylococ
22	12	27.9	9	2	Q48687	Q48687 lactococcus
23	12	27.9	9	2	Q93LE4	Q93le4 helicobacter
24	12	27.9	9	2	Q93E20	Q93e20 streptococc
25	12	27.9	9	4	Q9H4B0	Q9h4b0 homo sapien
26	12	27.9	9	6	Q28112	Q28112 bos taurus
27	12	27.9	9	15	Q70140	Q70140 human human
28	11	25.6	5	13	P82072	P82072 litoria rob
29	11	25.6	5	13	P82073	P82073 litoria rob
30	11	25.6	6	10	P82181	P82181 spinacia ol
31	11	25.6	6	10	P82182	P82182 spinacia ol
32	11	25.6	7	2	Q47505	Q47505 escherichia
33	11	25.6	8	2	Q9RFL7	Q9rfl7 clostridium
34	11	25.6	8	2	Q9S6D5	Q9s6d5 escherichia
35	11	25.6	8	2	Q9R7E2	Q9r7e2 escherichia
36	11	25.6	8	2	Q44463	Q44463 rhizobiaceа
37	11	25.6	8	3	P87225	P87225 saccharomyc
38	11	25.6	8	4	Q1b428	Q1b428 homo sapien
39	11	25.6	8	6	Q9MYL5	Q9myl5 pongo pygma
40	11	25.6	8	7	Q95213	Q95213 oryctolagus
41	11	25.6	8	7	Q29801	Q29801 homo sapien
42	11	25.6	8	7	Q29810	Q29810 homo sapien
43	11	25.6	8	7	Q29797	Q29797 homo sapien
44	11	25.6	8	7	Q29812	Q29812 homo sapien
45	11	25.6	8	8	Q9I2Y2	Q9i2y2 boquoria tri

ALIGNMENTS

RESULT 1
Q85710
ID Q85710 PRELIMINARY; PRT; 9 AA.
AC Q85710;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE POL PROTEIN (FRAGMENT).
OS Rous sarcoma virus.
OC Viruses; Retroid viruses; Retroviridae; Alpharetrovirus.
OX NCBI_TaxID=11886;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=84115080; PubMed=6319754;
RA Lerner T.L., Hanafusa H.;
RT "DNA sequence of the Bryan high-titer strain of Rous sarcoma virus:
RT gag-pol-src deletion and possible noncontinuous recombinational relationship with
RT other viral strains.";
RL J. Virol. 49:549-556(1984).
DR EMBL; K03365; AAA42557.1; -.
FT NON_TER 1 1
SQ SEQUENCE 9 AA; 949 MW; 94AA1441E0E07A1AA CRC64;

Query Match 48.8%; Score 21; DB 15; Length 9;
Best Local Similarity 57.1%; Pred. No. 5.9e+05;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 2 TLASSET 8
||:::|:
Db 3 TLAANES 9

RESULT 2
P82858
ID P82858 PRELIMINARY; PRT; 8 AA.
AC P82858;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)

```
DT    [illegible] (TrEMBLrel. 17, Last annotation update)
DE    URICASE (EC 1.7.3.3) (URATE OXIDASE) (FRAGMENT).
GN    [illegible].
OS    Candida boidinii [illegible] sp. trisoti.
OC    Eukaryota; Fungi; Basidiomycota; Hemiascomycetes; Candidomycetidae;
OC    Saccharomycetales; Pichiaceae; Pichia.
OX    NCBI_TaxID=34279;
RN    [1]
RP    SEQUENCE FROM N.A.
RC    [illegible] FUNCTION, AND DEVELOPMENTAL STAGE.
RC    [illegible] TISSUE;
RA    Achilli M., Montalbini P., Pileda M.;
RL    Submitted (NOV 2000) to the SWISS-PROT data bank.
CC    -!- FUNCTION: INVOLVED IN HOST-PARASITE RELATIONS BETWEEN PLANTS
CC        AND FUNGI.
CC    -!- CATALYTIC ACTIVITY: URATE + O(2) + H(2)O = 5-HYDROXYISOURATE +
CC        H(2)O(2) (2). (2) HYDROXYISOURATE DECOMPOSE TO FORM ALLANTOIN).
CC    -!- SUBCELLULAR LOCATION: PEROXISOMAL (BY SIMILARITY).
CC    -!- DEVELOPMENTAL STAGE: TERMINATE N.
CC    -!- SIMILARITY: BELONGS TO THE URICASE FAMILY.
DR    InterPro; IPR002042; Uricase.
DR    [illegible] SIZE; PS00366; URICASE; PARTIAL.
KW    [illegible]oxidoreductase; Purine metabolism; Peroxisome.
FT    NON_TER       8      8
SQ    SEQUENCE   8 AA;  777 MW;  98C1AD073B9B76D CRC64;


  Query Match             50.0%;  Score 27;  DB 3;  Length 8;
  Best Local Similarity   50.0%;  Pred. No. 5.6e+04;
  Matches    3;  Conservative    4;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 FILASS 6
           :  :::
Db       1 YIISSA 6


RESULT   4
[illegible]OMW4
ID   [illegible]OMW4       PRELIMINARY;      PRT;     9 AA.
AC   [illegible]OMW4;
DT   [illegible] (TrEMBLrel. 15, Created)
DT   [illegible] (TrEMBLrel. 15, Last sequence update)
DT   [illegible] (TrEMBLrel. 16, Last annotation update)
DE   MHC CLASS I ANTIGEN (FRAGMENT).
GN   HLA B4900.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=BLOOD;
RX   MEDLINE=97249738; PubMed=9089103;
RA   Varela-Calvino R., Gomez-Casado E., Martinez-Laso J., Granados J.,
RA   Lazaro A., Arnaiz-Villena A.;
RT   "Sequences in tuition 2 sequences between B*4900:1 and B*4902 in
RT   Amerindians: comparison with those of B*4901, B*4501, and B*4002
RT   alleles.";
RL   Immunogenetics 45:136-149(1997).
DR   EMBL; [illegible]; AAB90298.1; -.
FT   NON_TER       1      1
FT   NON_TER       9      9
SQ   SEQUENCE   9 AA;  975 MW;  51C661AB01A7213F CRC64;


  [illegible]Match             50.0%;  Score 27;  DB 2;  Length 9;
  Best Local Similarity   66.7%;  Pred. No. 5.6e+04;
  Matches    4;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 ASSEI 6
           :  :::
Db       [illegible]A:CEI 7


RESULT   5
```

```
P72221
ID   P72221       PRELIMINARY;      PRT;     8 AA.
AC   P72221;
DT   01-FEB-1997 (TrEMBLrel. 02, Created)
DT   01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   ALGINATE LYASE (FRAGMENT).
GN   ALY.
OS   Pseudomonas sp. (strain OS-ALG-9).
OC   Bacteria; Proteobacteria.
OX   NCBI_TaxID=86038;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=OS-ALG-9;
RX   MEDLINE=93329366; PubMed=8336113;
RA   Maki H., Mori A., Fujiyama K., Kinoshita S., Yoshida T.;
RT   "Cloning, sequence analysis and expression in Escherichia coli of a
RT   gene encoding an alginate lyase from Pseudomonas sp. OS-ALG-9.";
RL   J. Gen. Microbiol. 139:987-993(1993).
DR   EMBL; L08469, BAA21704.1; -.
KW   Lyase.
FT   NON_TER       1      1
FT   NON_TER       8      8
SQ   SEQUENCE   8 AA;  841 MW;  461DDDC5A5B041BB CRC64;


  Query Match             37.2%;  Score 16;  DB 2;  Length 8;
  Best Local Similarity   33.3%;  Pred. No. 5.6e+05;
  Matches    2;  Conservative    4;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 FILASS 6
           :  :::
Db       1 YIISSA 6


RESULT   6
Q60615
ID   Q60615       PRELIMINARY;      PRT;     8 AA.
AC   Q60615;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
DE   APOLIPOPROTEIN A-II (FRAGMENT).
GN   APOA2.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciuromorpha; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN C57BL/6J;
RX   MEDLINE=94319082; PubMed=8043949;
RA   Ko M.S., Wang X., Horton J.H., Hagen M.D., Takahashi N., Maeda Y.,
RA   Nadeau J.H.;
RT   "Genetic mapping of 40 cDNA clones on the mouse genome by FCS.";
RL   Mamm. Genome 5:349-355(1994).
DR   EMBL; U05691; AAB60462.1; -.
DR   MGD; MGI:88050; Apoa2.
KW   Lipoprotein.
FT   NON_TER       1      1
SQ   SEQUENCE   8 AA;  720 MW;  783D4DC5B861AB18 CRC64;


  Query Match             37.2%;  Score 16;  DB 11;  Length 8;
  Best Local Similarity   100.0%;  Pred. No. 5.6e+05;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       7 EIS 9
```

```
DR   [illegible] 4

RESULT   6
Q9ASQ0
ID   Q9ASQ0      PRELIMINARY;      PRT;     9 AA.
AC   Q9ASQ0;
DT   01-DEC-2000 (TrEMBLrel. 19, Created)
DT   01-DEC-2000 (TrEMBLrel. 19, Last sequence update)
DT   01-DEC-2000 (TrEMBLrel. 19, Last annotation update)
DE   HYPOTHETICAL 0.9 KDA PROTEIN (FRAGMENT).
GN   pCM2.60 1.
OS   Salmonella enterica subsp. enterica serovar Typhi.
OC   Plasmid pCM2.
OC   Bacteria; Proteobacteria; Gamma subdivision; Enterobacteriaceae;
OC   Salmonella.
OX   NCBI_TaxID=90370;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CT18;
RX   MEDLINE 21534947; PubMed=11677608;
RA   Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J.,
RA   Churcher C., Mungall K.L., Bentley S.D., Holden M.T.G., Sebaihia M.,
RA   Baker S., Basham D., Brooks K., Chillingworth T., Connerton P.,
RA   Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,
RA   Feltwell T., Hamlin N., Haque A., Hien T.T., Holroyd S., Jagels K.,
RA   Krogh A., Larsen T.S., Leather S., Moule S., O'Gaora P., Parry C.,
RA   Quail M., Rutherford K., Simmonds M., Skelton J., Stevens K.,
RA   Whitehead S., Barrell B.G.;
RT   "Complete genome sequence of a multiple drug resistant Salmonella
RT   enterica serovar Typhi CT18.";
RL   Nature 413:848-852(2001).
DR   EMBL; AL513384; CAD09868.1; -.
KW   Hypothetical protein; Plasmid.
FT   NON_TER       9       9
SQ   SEQUENCE   9 AA;  984 MW;  [illegible] CRC64;


  Query Match             47.2%;  Score 16;  DB 2;  Length 9;
  Best Local Similarity   66.7%;  Pred. No. 5.6e+05;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       2 ILASSE 7
           ||:  |
Db       4 IIAGIE 9


RESULT   7
Q9XSL0
ID   Q9XSL0      PRELIMINARY;      PRT;     9 AA.
AC   Q9XSL0;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   ALPHA S2-CASEIN (FRAGMENT).
GN   CSN1S2.
OS   Capra hircus (Goat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC   Caprinae; Capra.
OX   NCBI_TaxID=9925;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Blood;
RX   MEDLINE 21419340; PubMed=11419340;
RA   Ramunno L., Longobardi E., Pappalardo M., Rando A., Di Gregorio F.,
RA   Cosenza G., Mariani P., Pastore N., Masina P.;
RT   "An allele associated with a non detectable amount of as2 casein in
RT   goat milk.";
RL   Anim. Genet. 32:19-26(2001).
DR   EMBL; AJ238475; CAB44298.1; -.
FT   NON_TER       1       1
FT   VARIANT       5       5       I -> V.
FT   NON_TER       9       9
SQ   SEQUENCE   9 AA;  1047 MW;  AAA2A5A5B0F1EE1D CRC64;


  Query Match             47.2%;  Score 16;  DB 6;  Length 9;
  Best Local Similarity   60.0%;  Pred. No. 5.6e+05;
  Matches    3;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       4 LASSE 7
           :::|||
Db       5 ISSSE 9


RESULT   8
Q9QVH9
ID   Q9QVH9      PRELIMINARY;      PRT;     9 AA.
AC   Q9QVH9;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   SUPERINDUCIBLE PROTEIN 24, SIP24-CYCLOPHILIN HOMOLOG, PEAK C.
OS   Mus sp.
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciuromorpha; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10095;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE 92031730; PubMed=1933154;
RA   Davis T.R., Tabatabai L., Prigge R., Barbera R.L., Nelson Rambosek M.;
RT   "Basic fibroblast growth factor induces its fibroblasts to synthesize
RT   and secrete a cyclophilin like protein and beta 2 microglobulin.";
RL   Biochim. Biophys. Acta 1095:145-152(1991).
SQ   SEQUENCE   9 AA;  963 MW;  D612273A1D029CD CRC64


  Query Match             47.2%;  Score 16;  DB 11;  Length 9;
  Best Local Similarity   60.0%;  Pred. No. 5.6e+05;
  Matches    3;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 FTLAS 5
           ||| :
Db       5 FILGN 9


RESULT   9
Q9QZA8
ID   Q9QZA8      PRELIMINARY;      PRT;     9 AA.
AC   Q9QZA8;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   C-TYPE LECTIN DCLI (FRAGMENT).
GN   DCLI.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciuromorpha; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Gorski K., Boang X., Iseki S.-Y., Kattis E., Parrott ., Ishchiya H.;
RT   "Dendritic cell population of DCLI mRNA-expressing.";
RL   Submitted (OCT-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF192526; AAF04843.1; -.
DR   MGD; MGI:2136650; Dcli.
KW   Lectin.
FT   NON_TER       9       9
SQ   SEQUENCE   9 AA;  994 MW;  84216AB17AFBAB7 CRC64;


  Query Match             47.2%;  Score 16;  DB 11;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 5.6e+05;
  Matches    3;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

Qy 7 EEE 9

Db 6 EEE 8

RESULT 11
P76864
ID P76864 PRELIMINARY; PRT; 7 AA.
AC P76864;
DT 01-FEB-1997 (TrEMBLrel. 02, Created)
DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT 15-DEC-2000 (TrEMBLrel. 19, Last annotation update)
DE ATP SYNTHASE (FRAGMENT).
GN ATPG.
OS Azotobacter vinelandii.
OC Bacteria; Proteobacteria; Gamma subdivision; Pseudomonadales;
OC Azotobacter.
OX NCBI_TaxID=354;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN O;
RX MEDLINE=97427418; PubMed=8863082;
RA Poole R.K.A., Eftesvaq H., Valla S.;
RT "A new Azotobacter vinelandii mannuronan C-5-epimerase-gene (algG) is
RT part of an algE-like gene cluster physically organized in a manner similar
RT to that in Pseudomonas aeruginosa.";
RL Microbiology 142:1085-1092(1996).
DR EMBL; X83873; CAA61230.1; -.
FT NON_TER 1 1
SQ SEQUENCE 7 AA; 684 MW; 7185A5A5A2D1AE0C CRC64;

Query Match 34.9%; Score 15; DB 2; Length 7;
Best Local Similarity 60.0%; Pred. No. 5.6e+05;
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 5 SSEEG 9
| |:||
Db 2 SREEG 6

RESULT 12
Q94623
ID Q94623 PRELIMINARY; PRT; 8 AA.
AC Q94623;
DT 01-FEB-1997 (TrEMBLrel. 02, Created)
DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE MSUSP-2 PROTEIN (FRAGMENT).
GN USP.
OS Manduca sexta (Tobacco hawkmoth) (Tobacco hornworm).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
OC Sphingoidea; Sphingidae; Sphinginae; Manduca.
OX NCBI_TaxID=7130;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=EPIDERMIS, WING;
RX MEDLINE=97165493; PubMed=9018254;
RA Jindra M., Huang J.Y., Malone F., Asahina M., Riddiford L.M.;
RT "Identification and mRNA developmental profiles of two ultraspiracle
RT isoforms in the epidermis and wings of Manduca sexta.";
RL Insect Mol. Biol. 6:41-53(1997).
DR EMBL; U57921; AAB64235.1; -.
FT NON_TER 8 8
SQ SEQUENCE 8 AA; 892 MW; F165BB0415A76B16 CRC64;

Query Match 34.9%; Score 15; DB 5; Length 8;
Best Local Similarity 60.0%; Pred. No. 5.6e+05;
Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 5 SSEEG 9
| |:||
Db 4 SRESG 8

RESULT 13
Q66545
ID Q66545 PRELIMINARY; PRT; 9 AA.
AC Q66545;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE EPITOPE #14 (FRAGMENT).
OS Human herpesvirus 4 (Epstein-Barr virus).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Gammaherpesvirinae; Lymphocryptovirus.
OX NCBI_TaxID=10376;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN B95-8;
RX MEDLINE=88296424; PubMed=2841116;
RA Walls D., Gannon F.;
RT "The expression of novel antigens from the Epstein-Barr virus latent
RT internal repeat.";
RL EMBO J. 7:1191-1196(1988).
DR EMBL; X07817; CAA30676.1; -.
FT NON_TER 1 1
FT NON_TER 9 9
SQ SEQUENCE 9 AA; 929 MW; 82740B1DD735B735 CRC64;

Query Match 34.9%; Score 15; DB 12; Length 9;
Best Local Similarity 57.5%; Pred. No. 5.6e+05;
Matches 3; Conservative 3; Mismatches 2; Indels 1; Gaps 0;

Qy 2 TSASSEIG 9
:|: :| |
Db 2 SLSLAERG 9

```
RESULT  14
Q84977
ID   Q84977      PRELIMINARY;      PRT;     8 AA.
AC   Q84977;
DT   01-NOV-1996 (IrEMBLrel. 01, Created)
DT   01-NOV-1996 (IrEMBLrel. 01, Last sequence update)
DT   01-DEC-2001 (IrEMBLrel. 19, Last annotation update)
DE   INTERNAL A/[illegible] (H3N2). NUCLEOTIDES ([illegible]
DE   (FRAGMENT).
OS   Influenza A virus.
OC   Viruses; ssRNA negative-strand viruses; orthomyxoviridae;
OC   Influenza virus A and B group; Influenza A viruses.
OX   NCBI_TaxID=11320;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE 84112211; PubMed 6296449;
RA   Lin B.-Y., Lai C.-J.;
RT   "The influenza virus nucleoprotein synthesized from cloned DNA in a
RT   simian virus 40 vector is detected in the nucleus.";
RL   J. Virol. 45:434-438(1983).
DR   EMBL; J02176; AAA43468.1; -.
KW   Nucleocapsid.
FT   NON_TER       8       8
SQ   SEQUENCE   8 AA;  878 MW;  ED0421A8660C5BD16 CRC64;


  Query Match             80.2%;  Score 14;  DB 1;  Length 8;
  Best Local Similarity   50.0%;  Pred. No. 5.6e+05;
  Matches    3;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       3 LASSET 8
           :
Db       1 MASGII 6

RESULT  15
Q9VV82
ID   Q9VV82      PRELIMINARY;      PRT;     9 AA.
AC   Q9VV82;
DT   01-MAY-2000 (IrEMBLrel. 13, Created)
DT   01-MAY-2000 (IrEMBLrel. 13, Last sequence update)
DT   01-MAY-2000 (IrEMBLrel. 13, Last annotation update)
DE   CG18219 PROTEIN (FRAGMENT).
GN   CG18219.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN BERKELEY;
RX   MEDLINE 20166096; PubMed 10731132;
RA   Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA   Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA   George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA   Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA   Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA   Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA   Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA   Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA   Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA   Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA   Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA   Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA   de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA   Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA   Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA   Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA   Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA   Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA   Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA   Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA   Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA   Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA   Liu X., Mattei B., McIntosh T.C., Mcleod M.P., McPherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA   Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA   Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT   "The genome sequence of Drosophila melanogaster.";
RL   Science 287:2185-2195(2000).
DR   EMBL; AE003526; AAF49437.1; -.
DR   FlyBase; FBgn0036647; CG18219.
FT   NON_TER       1       1
SQ   SEQUENCE   9 AA;  975 MW;  F31AB447234516C7 CRC64;


  Query Match             80.2%;  Score 14;  DB 5;  Length 9;
  Best Local Similarity   44.4%;  Pred. No. 5.6e+05;
  Matches    2;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 FTLASS 6
           |:: :|
Db       4 FSILNS 8


Search completed: August 26, 2002, 06:39:45
Job time: 976 sec
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

Mprotein protein search, using sw model

Run on: August 26, 2002, 06:27:08 ; Search time 1.0611 Seconds
(without alignments)
7.903 Million cell updates/sec

Title: US-09-744-282-2
Perfect score: 48
Sequence: 1 FILASSEIG 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 1.49e+

Minimum DB seq length: 0
Maximum DB seq length: 9

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_082802:*
1: /SIDS5/data/geneseq/geneseq-ip/ebi_001.seq:* *
2: /SIDS5/data/geneseq/geneseq-ip/embl_AA1.seq:* *
3: /SIDS5/data/geneseq/geneseq-ip/ebi_002.seq:* *
4: /SIDS5/data/geneseq/geneseq-ip/embl_AA2.seq:* *
5: /SIDS5/data/geneseq/geneseq-ip/embl_AA3.seq:* *
6: /SIDS5/data/geneseq/geneseq-ip/embl_AA4.seq:* *
7: /SIDS5/data/geneseq/geneseq-ip/embl_AA5.seq:* *
8: /SIDS5/data/geneseq/geneseq-ip/ebi_003.seq:* *
9: /SIDS5/data/geneseq/geneseq-ip/embl_AA6.seq:* *
10: /SIDS5/data/geneseq/geneseq-ip/ebi_004.seq:* *
11: /SIDS5/data/geneseq/geneseq-ip/ebi_005.seq:* *
12: /SIDS5/data/geneseq/geneseq-ip/ebi_006.seq:* *
13: /SIDS5/data/geneseq/geneseq-ip/ebi_007.seq:* *
14: /SIDS5/data/geneseq/geneseq-ip/ebi_008.seq:* *
15: /SIDS5/data/geneseq/geneseq-ip/ebi_009.seq:* *
16: /SIDS5/data/geneseq/geneseq-ip/ebi_010.seq:* *
17: /SIDS5/data/geneseq/geneseq-ip/ebi_011.seq:* *
18: /SIDS5/data/geneseq/geneseq-ip/ebi_012.seq:* *
19: /SIDS5/data/geneseq/geneseq-ip/ebi_013.seq:* *
20: /SIDS5/data/geneseq/geneseq-ip/ebi_014.seq:* *
21: /SIDS5/data/geneseq/geneseq-ip/ebi_015.seq:* *
22: /SIDS5/data/geneseq/geneseq-ip/ebi_016.seq:* *

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	48	100.0	9	17	AAR95072	Human cartilage gl
2	48	100.0	9	21	AAY69227	Peptide fragment o
3	34	70.1	9	21	AAY77399	Human chondrocyte
4	24	50.0	6	8	AAP71386	Diuretic hexapepti
5	23	47.9	8	18	AAW22885	Elastin-derived
6	23	47.9	9	17	AAR95071	Human cartilage gl
7	23	47.9	9	21	AAY69226	Peptide fragment o
8	21	43.8	7	12	AAR95517	Human prostate car
9	21	43.8	8	21	AAY64587	Cadherin related n
10	21	43.8	9	14	AAR33422	Rat VL-12 sequence
11	21	43.8	9	16	AAR87292	Plasmodium falcipa
12	21	43.8	9	18	AAW85522	Rat T cell recepto
13	21	43.8	9	18	AAW17496	Protein kinase C m
14	21	43.8	9	21	AAY64546	Cadherin related n
15	21	43.8	9	22	AAG88904	HER2/neu epitope B
16	21	43.8	9	22	AAU00550	Human MUC1 peptide
17	21	43.8	9	22	AAB76129	Tumour associated
18	21	43.8	9	22	AAB76130	Tumour associated
19	20	41.7	5	21	AAY60996	Cadherin 5 cell ad
20	20	41.7	5	21	AAY61333	Cadherin 5 cell ad
21	20	41.7	6	21	AAY60998	Cadherin 5 cell ad
22	20	41.7	6	21	AAY61335	Cadherin 5 cell ad
23	20	41.7	7	15	AAR46056	HLA restricted T h
24	20	41.7	7	16	AAR72778	Mammalian ribosoml
25	20	41.7	7	21	AAY88856	Ribonucleotide red
26	20	41.7	7	21	AAY60996	Cadherin 5 cell ad
27	20	41.7	7	21	AAY61120	Cadherin 5 cell ad
28	20	41.7	7	21	AAY61264	Cadherin 5 cell ad
29	20	41.7	7	21	AAY61208	Cadherin 5 cell ad
30	20	41.7	7	21	AAY61256	Cadherin 5 cell ad
31	20	41.7	7	21	AAY61296	Cadherin 5 cell ad
32	20	41.7	7	21	AAY61338	Cadherin 5 cell ad
33	20	41.7	8	16	AAR72456	Human ISH recept a
34	20	41.7	8	21	AAY60999	Cadherin 5 cell ad
35	20	41.7	8	21	AAY61123	Cadherin 5 cell ad
36	20	41.7	8	21	AAY61187	Cadherin 5 cell ad
37	20	41.7	8	21	AAY61211	Cadherin 5 cell ad
38	20	41.7	8	21	AAY61255	Cadherin 5 cell ad
39	20	41.7	8	21	AAY61298	Cadherin 5 cell ad
40	20	41.7	8	21	AAY61341	Cadherin 5 cell ad
41	20	41.7	8	21	AAY64432	Cadherin related n
42	20	41.7	8	22	AAB75579	[illegible]
43	20	41.7	8	22	AAB77[illegible]	[illegible]
44	20	41.7	8	22	AAP[illegible]	[illegible] terminal co
45	20	41.7	9	19	AAW70079	S. cerevisiae meth

ALIGNMENTS

RESULT 1
AAR95072
ID AAR95072, standard; peptide; 9 AA.
XX
AC AAR95072;
XX
DT 17-SEP-1996 (first entry)
XX
DE Human cartilage glycoprotein-39 autoantigen derived peptide (11).
XX
KW Human; cartilage; glycoprotein-39; hcgp-39; autoantigen;
KW tolerization; T cell; disease; modulator; destruct ion;
KW rheumatoid arthritis; disease model; diagnosis.
XX
OS Homo sapiens.
XX
PN WO9614517-A1.
XX
PD 09-MAY-1996.
XX
PF 25-OCT-1995; 95WO-EP04201.
XX
PR 07-APR-1995; 95EP-0200888.
PR 27-OCT-1994; 94EP-0203128.
XX
PA (AKZO) AKZO NOBEL NV.
XX
PI Boots AMH, Verheijden GFM;
XX
DR WPI; 1996-23944/24.
XX
PT Peptide(s) derived from human cartilage gp-39 and variants used
PT to induce specific T cell tolerance, esp. for treatment of
PT rheumatoid arthritis, also useful to produce cells in animals.

```
RC   Residues
XX
RN   [1]
RP   1-9 FEB 2000; 3pp; English.
XX
RA   "Peptides derived from the human cartilage glycoprotein
RA   39 (hc gp 39) autoantigen, can be used to induce specific T-cell
RA   tolerance to hc gp 39 in a patient with T-cell mediated cartilage
RA   destruction, esp. rheumatoid arthritis (RA). It may also be used to
RA   reduce RA in mice, etc. (human disease models), and to diagnose the
RA   presence of activated T cells autoreactive with hc gp 39, i.e. to
RA   identify patients suitable for tolerising treatment.
RA   Preferably the peptide is given in a dose of 0.1 to microg/kg
RA   by injection, or 10-50 microg/kg to produce a disease model. The
RA   peptide has a specific tolerising effect on autoreactive T cells,
RA   but leaves other components of the immune system intact, and
RA   therefore should have low side effects. In an assay to determine
RA   the peptide's binding to HLA DR molds., specificity shown (Kd=xx
RA   M for associated with an increased risk of RA development), the
RA   peptide gave a T-1 plot IC(50) (micromolar) value of 0.00 compared
RA   0.2 for the control peptide TRA(307-319)".
XX
SQ   Sequence    9 AA;


  Query Match             100.0%;  Score 43;  DB 21;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 6.4e-05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        1 FTLASSETG 9
            |||||||||
Db        1 FTLASSETG 9


RESULT    3
AAY77399
ID   AAY77399 standard; peptide; 9 AA.
XX
AC   AAY77399;
XX
DT   22-MAY-2000  (first entry)
XX
DE   Human chondrocyte protein YKL-39, residues 268-276.
XX
KW   Human chondrocyte protein; YKL 39; epitope; autoreactive T cell;
KW   tolerisation; desensitisation; autoantigen; rheumatoid arthritis;
KW   HC gp-39; human cartilage glycoprotein-39; YKL-40; bystander suppression;
KW   autoimmune disorder; Graves' disease; primary glomerulonephritis;
KW   inflammatory bowel disease; diagnosis.
XX
OS   Homo sapiens.
XX
PN   WO200005254-A2.
XX
PD   03-FEB-2000.
XX
PF   16-JUL-1999;   99WO-EP005050.
XX
PR   23-JUL-1998;   98EP-0202470.
XX
PA   (ALKU ) AKZO NOBEL NV.
XX
PI   Verheijden GFM,  Boots AMH;
XX
DR   WPI; 2000-182644/16.
XX
PT   New T-cell epitope peptides useful for inducing specific T-cell
PT   tolerance to an autoantigen for treating autoimmune disorders,
PT   especially rheumatoid arthritis -
XX
PS   Claim 1; Page 16; 26pp; English.
XX
CC   Sequences AAY77399-Y77400 respectively represent amin. acids 259-276 and
CC   amino acids 266-278 of human chondrocyte protein YKL-39. The peptides
CC   may be used in pharmaceutical compositions for the induction of specific
CC   T-cell tolerance (desensitisation) to an autoantigen in patients
CC   suffering from autoimmune disorders, especially rheumatoid arthritis. A
CC   frequently found autoantigen in rheumatoid arthritis sufferers is human
CC   cartilage glycoprotein 39 (HC gp 39, also known as YKL 40). Several
CC   epitopes of HC gp 39 are recognised by peripheral blood T cells from
CC   rheumatoid arthritis patients; one such epitope, HC gp 39 residues
CC   263-275, was recognised by T cells from 8 out of 18 patients tested.
CC   YKL 39 shares significant sequence identity to hc gp 39. YKL 39 residues
CC   266-278 (AAY77400) constitutes a mimicry epitope of HC gp 39 (263-275),
CC   and may be used for desensitisation of T cells with reactivity to HC
CC   gp 39 (263-275) in rheumatoid arthritis patients. Rheumatoid arthritis
```

```
CC   may be treated by the systemic administration of compositions comprising
CC   the peptides of the invention, optionally in combination with existing
CC   treatments. The peptides may also be used for modulating lymphocytes
CC   that are reactive to autoantigens other than YKL 39 and HC gp 39 but
CC   which are present in the same tissue as these autoantigens, by the
CC   induction of antigen specific T-cell tolerance. Autoimmune disorders can
CC   be treated by bystander suppression. The conditions that may be treated
CC   via this approach include Graves' disease, primary glomerulonephritis
CC   and inflammatory bowel disease. The peptides, alone or in a detection
CC   agent, may additionally be used in a diagnostic composition. The peptides
CC   of the invention have a specific effect on autoreactive T cells whilst
CC   leaving other components of the immune system intact. Therefore, the
CC   peptides should be safe, having no side effects.
XX
SQ   Sequence    9 AA;


  Query Match             79.1%;  Score 34;  DB 21;  Length 9;
  Best Local Similarity   87.5%;  Prod. No. 6.4e-05;
  Matches    7;  Conservative    1;  Mismatches    0;  Indels    1;  Gaps    0;

Qy       1 PIIASSET 8
           |:|:|:|.|
Db       1 PTIASDET 8


RESULT   4
AAP71386
ID   AAP71386 standard; peptide; 6 AA.
XX
AC   AAP71386;
XX
DT   07-MAY-1987  (first entry)
XX
DE   Diuretic hexapeptide.
XX
KW   Diuretic; natriuretic.
XX
OS   synthetic.
XX
FH   Key             Location/Qualifiers
FT   Peptide         1..2
FT                   /label= active fragment
FT   Peptide         3..4
FT                   /label= active fragment
FT   Peptide         5..6
FT                   /label= active fragment
XX
PN   FR2597486-A.
XX
PD   23-OCT-1987.
XX
PF   18-APR-1986;   86FR-0005584.
XX
PR   18-APR-1986;   86FR-0005584.
XX
PA   (FLOR/) FLORK M L A.
XX
PI   Flork MLA  Cotonat JN;
XX
DR   WPI; 1987-350400/50.
XX
PT   New hexapeptide and diuretic composi... -derived hexapeptide a
PT   its dipeptide fragments
XX
PS   Claim 1; page 10; 14pp; French.
XX
CC   This hexapeptide is synthesised using standard solid phase methods.
CC   Its sequence is based on that of a natural hexapeptide extracted
CC   from the lymph nodes of dogs. This hexapeptide and its dipeptide
CC   fragments are useful as diuretics and natriuretics.
XX
SQ   Sequence    6 AA;


  Query Match             55.8%;  Score 24;  DB 8;  Length 6;
  Best Local Similarity   83.3%;  Prod. No. 7.4e-05;
  Matches    5;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       4 ASSEIG 9
           |:|:||
Db       1 asdetq 6


RESULT   5
AAW22885
ID   AAW22885 standard; peptide; 8 AA.
XX
AC   AAW22885;
XX
DT   25-SEP-1997  (first entry)
XX
DE   Plasminogen activator inhibitor 1, inactivating peptide.
XX
KW   Plasminogen activator inhibitor; PAI 1; inactivating; thrombosis;
KW   clot lysis; half life; reduction; vitronectin; protease;
KW   tissue plasminogen activator; urokinase; plasminogen; inhibition;
KW   activation.
XX
OS   Homo sapiens.
XX
PN   US5639726-A.
XX
PD   17-JUN-1997.
XX
PF   30-SEP-1994;   94US-0315461.
XX
PR   30-SEP-1994;   94US-0315461.
XX
PA   (FORD ) FORD HEALTH SYSTEM HENRY.
PA   (UNMI ) UNIV MICHIGAN.
XX
PI   Eitzman DT,  Fay WP,  Francis Chmura AM,  Ginsburg D;
PI   Lawrence DA,  Olson ST,  Palelli D,  Shore JD;
XX
DR   WPI; 1997-332067/30.
XX
PT   Peptide(s) that inactivate plasminogen activator inhibitor 1
PT   useful for promoting clot lysis
XX
PS   Claim 1; Columns 77-78; 54pp; English.
XX
CC   The present peptide is a plasminogen activator inhibitor 1 (PAI 1)
CC   inactivating peptide, which can promote clot lysis by decreasing
CC   the half life of PAI-1, preferably to 1-90 (especially 1-10)
CC   minutes in vitro or to 1-190 (especially 110-130) minutes in vitro
CC   in the presence of vitronectin. The peptide promotes the conversion
CC   of PAI-1 from its active state to an inactive state and thus
CC   renders it incapable of complexing with its target proteases,
CC   namely tissue plasminogen activator and urokinase.
CC   The present peptide comprises residues 356 to 363 of PAI 1.
XX
SQ   Sequence    8 AA;


  Query Match             53.5%;  Score 23;  DB 8;  Length 8;
  Best Local Similarity   71.4%;  Prod. No. 7.4e-05;
  Matches    5;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       2 TIASSET 8
           |:|||.|
Db       1 tvassst 7


RESULT   6
AAR95071
```

```
ID   AAR[illegible] standard; peptide; [illegible] AA.
XX
AC   AAR[illegible];
XX
DT   [illegible] (first entry)
XX
DE   Bit[illegible] cartilage glycoprotein 39 autoantigen derived peptide (1).
XX
KW   [illegible] cartilage glycoprotein 39; hC gp 39; rheumatoid;
KW   arthritis; T cell tolerance; mediator; destructive;
KW   RA and OA arthritis; disease model; diagnosis.
XX
OS   Homo sapiens.
XX
PN   WO[illegible]-A1.
XX
PD   [illegible] MAY 1999.
XX
PF   [illegible] NOV 1998;   98WO-EP07201.
XX
PR       AUG 1997;   97EP-0202996.
PR   [illegible] NOV 1997;   97EP-0204128.
XX
PA   (ALKU ) AKZO NOBEL NV.
XX
PI   B[illegible] AMH;  Verbeijden JFM;
XX
DR   WPI; [illegible]-447724.
XX
PT   [illegible]-B(1) derived from human cartilage gp 39 autoantigen, used
PT   [illegible] specific T cell tolerance, esp. for treatment of
PT   rheumatoid arthritis, also useful in predicting arthritis onset
PT   [illegible]
XX
PS   Claim 1; Page 20; 37pp; English.
XX
CC   The present peptide, derived from the human cartilage glycoprotein
CC   39 (hC gp 39) autoantigen, can be used to induce specific T cell
CC   tolerance to hC gp 39 in a patient with T cell mediated cartilage
CC   destruction, esp. rheumatoid arthritis (RA). It may also be used to
CC   induce RA in mice, etc. (human disease models), and to diagnose the
CC   presence of activated T cells autoreactive with hC gp 39, i.e. to
CC   identify patients suitable for tolerising treatment
CC   Therapeutically, the peptide is given in a dose of [illegible] (preferably
CC   between [illegible] to 50 mg/treatment/kg to produce a disease model. The
CC   peptide has a specific tolerising effect on autoreactive T cells,
CC   i.e. causes effector components of the immune system to [illegible] and
CC   therefore should have few side effects. In an assay to determine
CC   the peptide's binding to HLA DR mols., specifically DR1w4 (known
CC   to be associated with an increased risk of RA-developing), the
CC   peptide gave a 50 % [illegible] (micromolar) value of [illegible] compared
CC   to [illegible] for the control peptide TBA(307-319)F.
XX
SQ   Sequence  [illegible] AA;
```

```
  Query Match             [illegible]%;  Score 24;  DB 17;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. [illegible];
  Matches    [illegible];  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       [illegible] FTLAS 5
         [illegible]
Db       [illegible] FTLAS 9
```

```
RESULT   [illegible]
AAR[illegible]
ID   AAR[illegible] standard; peptide; [illegible] AA.
XX
AC   AAR[illegible];
XX
DT   [illegible] (first entry)
XX
```

```
ID   AAR95517 standard; peptide; 7 AA.
XX
AC   AAR95517;
XX
DT   06-NOV-1996  (first entry)
XX
DE   Human prostate carcinoma cell antigen binding peptide ("abtide").
XX
KW   Abtide; prostate specific mucin antigen; human; prostate cancer; PNCa1;
KW   diagnostic; detection; imaging; tumour; phage; peptide library;
KW   polymorphic; epithelial.
XX
OS   Synthetic.
XX
PN   WO9609411-A1.
XX
PD   28-MAR-1996.
XX
```

```
DE   Peptide fragment of human cartilage gp 39 (HC gp 39) protein.
XX
KW   Human cartilage gp 39 protein; HC gp 39 protein; T cell activator;
KW   inflammatory disease; lymphocyte; antigen; bystander suppression;
KW   antigen specific T-cell tolerance; autoimmune disorder;
KW   inflammatory disease; rheumatoid arthritis;
KW   autoimmune inflammatory disease; autoantigen.
XX
OS   Homo sapiens.
XX
PN   WO2000004917-A2.
XX
PD   03-FEB-2000.
XX
PF   19-JUL-1999;   99WO-EP05341.
XX
PR   23-JUL-1998;   98EP-0202471.
XX
PA   (AKZO ) AKZO NOBEL NV.
XX
PI   Miltenburg AMM,  Boots AMH;
XX
DR   WPI; 2000-182547/16.
XX
PT   Use of cartilage protein or peptides for treating inflammatory diseases
PT   especially rheumatoid arthritis -
XX
PS   Claim 2; Page 17; 29pp; English.
XX
CC   AAY69220-33 represent fragments of the human cartilage gp 39 (HC gp 39)
CC   protein. The peptides are T-cell activators. The peptides are used for
CC   preventing inflammatory diseases, by modulating lymphocytes reactive
CC   to antigens other than HC gp 39, present in the same tissue of
CC   HC gp 39. By the induction of antigen specific T cell tolerance,
CC   autoimmune disorders can be treated by bystander suppression. The
CC   HC gp 39 peptides are useful for preventing inflammatory diseases
CC   especially rheumatoid arthritis. The fragments may also be used for
CC   treating autoimmune inflammatory diseases. The HC gp 39 peptides are
CC   especially beneficial in autoimmune conditions in which the
CC   autoantigens associated with the condition are unknown.
XX
SQ   Sequence   9 AA;
```

```
  Query Match             53.5%;  Score 23;  DB 21;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 9.4e+05;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 FTLAS 5
           |||||
Db       5 FTLAS 9
```

PF 25-SEP-1995; 95WO-US11644.
XX
PR 27-JAN-1995; 95US-0388161.
PR 21-SEP-1994; 94US-0310192.
XX
PA (CYTI-) CYTOGEN CORP.
XX
PI Alfano C;
XX
DR WPI; 1996-184471/19.
XX
PT New isolated peptide(s) with specific binding activities - obtd. by
PT screening random peptide libraries, for use in diagnostic and
PT therapeutic compsns.
XX
PS Claim 48; Page 94; 106pp; English.
XX
CC AAB65511-B65520 are antigen binding peptides ("abtides"), which bind to
CC a human prostate carcinoma cell antigen. The abtides are identified
CC from random peptide libraries using specific ligand binding. Abtides
CC mimic the binding specificity of large molecules such as antibodies
CC and receptors but have a much smaller size allowing their production
CC at a lower cost and reducing the extent of their immunogenicity aiding
CC in vivo delivery. The abtides are useful for the diagnosis, detection,
CC imaging and treatment of disease, e.g. tumours, prostate cancer.
XX
SQ Sequence 7 AA;

Query Match 48.8%; Score 21; DB 17; Length 7;
Best Local Similarity 57.1%; Pred. No. 6.4e+05;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 5 LASSETG 9
| : | :
Db 1 IVSTNSG 7

RESULT 9
AAY64547
ID AAY64547 standard; Peptide; 8 AA.
XX
AC AAY64547;
XX
DT 02-MAR 2000 (first entry)
XX
DE Cadherin related neuronal receptor CAR cyclic peptide SEQ ID NO:8054.
XX
KW Modification; nonclassical cadherin mediated cell adhesion; CAR;
KW inhibition; cadherin extracellular domain; cell adhesion recognition;
KW OB cadherin; cadherin 5; cadherin-6; cadherin-7; cadherin 8;
KW cadherin-12; cadherin 14; cadherin-15; T-cadherin; FB cadherin;
KW cadherin related neuronal receptor; EI cadherin; protocadherin;
KW desmoglein; desmocollin; calcium binding; cancer; tumour; obesity;
KW rheumatoid arthritis; multiple sclerosis; diabetes; metastasis;
KW neurological disease; cyclic.
XX
OS Synthetic
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Modified-site 1..8
FT /note= "the terminal residues are cysteine residues which
FT other to form a cyclic peptide"
XX
PN WO9957144-A2.
XX
PD 11-NOV-1999.
XX
PF 5-MAY-1999; 99WO-CA000363.
XX
PR 06-MAY-1998; 98US-0073041.
PR 09-NOV-1998; 98US-0107859.
PR 20-JAN-1999; 99US-0234395.
PR 08-MAR-1999; 99US-0264516.
XX
PA (ADHE-) ADHEREX TECHNOLOGIES INC.
XX
PI Blaschuk OW, Gour BJ, Byers S;
XX
DR WPI; 2000-038791/03.
XX
PT New cadherin modulating agents, used for modulating nonclassical
PT cadherin-mediated functions for treating e.g. cancers, obesity,
PT rheumatoid arthritis, multiple sclerosis, diabetes or a neurological
PT disease.
XX
PS Claim 102; Page 220; 252pp; English.
XX
CC The present invention describes cadherin modulating agents (MA)
CC comprising peptides which comprise a nonclassical cadherin cell adhesion
CC recognition (CAR) sequence. The MAs can be used to: modulating
CC nonclassical cadherin-mediated functions. They can be used for e.g.
CC inhibiting adhesion of nonclassical cadherin expressing cells in a
CC mammal, enhancing delivery of a drug through the skin of a mammal,
CC enhancing delivery of a drug to a tumour in a mammal, treating cancer in
CC a mammal, inhibiting metastasis of a cancer in a mammal, inhibiting
CC angiogenesis in a mammal, inducing apoptosis in a nonclassical cadherin
CC expressing cell, preventing or treating obesity in a mammal, stimulating
CC blood vessel regression in a mammal, enhancing drug delivery to the
CC central nervous system, treating a demyelinating neurological disease,
CC increasing vasopermeability in a mammal, enhancing adhesion of
CC nonclassical cadherin-expressing cells, inhibiting synaptic stability in
CC a mammal, or preventing pregnancy in a mammal. The MAs can also be used to
CC e.g. enhancing or directing neurite outgrowth, facilitating wound
CC healing or reducing scar tissue, or enhancing adhesion of fetoplacental tissue
CC in a mammal. They can also be used for treating e.g. psoriasis,
CC arthritis, age related macular degeneration, multiple sclerosis and
CC diabetes. The products can also be used for detection and diagnosis and
CC in bioreactors. AAY60962 to AAY64572 represent specifically claimed
CC peptides, and AAY64573 to AAY64643 and AAY83183 to AAY83186 represent
CC sequences used in the exemplification of the present invention.
XX
SQ Sequence 8 AA;

Query Match 48.8%; Score 21; DB 21; Length 8;
Best Local Similarity 57.1%; Pred. No. 6.4e+05;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 5 LASSETG 9
| : | :
Db 2 IVSTNSG 8

RESULT 10
AAR83422
ID AAR83422 standard; peptide; 9 AA.
XX
AC AAR83422;
XX
DT 26-JUL 1993 (first entry)
XX
DE Rat VDJ12 sequence.
XX
KW TCR; T cell receptor; autoimmune disease; rheumatoid arthritis; BA;
KW J beta domain; V beta domain; T cell mediated autoimmune disease;
KW antagonists.
XX
OS Rattus rattus.
XX
PN WO9306145-A.
XX
PD 01-APR 1993.
XX
PF 23-SEP-1992; 92WO-US08094.

```
XX
PN   EP 0 [illegible] [illegible]
PD   24-OCT-[illegible]
PF   24-APR-[illegible]
XX
PA   (PEPR-) PEPRESEARCH INC.
XX
PI   Arnon R,
XX
DR   WPI; [illegible]/14.
XX
PT   Synthetic peptide or multiepitopic polypeptide(s) useful in the
PT   therapeutic and treatment of auto immune disorders, particularly
PT   rheumatoid arthritis
XX
PS   [illegible] Page 9; 51pp; English
XX
CC   The sequence is the rat VEGF2 sequence and is useful within a DNA
CC   or RNA molecule polypeptide.
XX
SQ   Sequence   9 AA;


  Query Match             48.8%;  Score 21;  DB 14;  Length 9;
  Best Local Similarity   66.7%;  Pred. No. 6.4e+05;
  Matches    4;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       3 ACSETI 8
           |:|
Db       3 AGSEAI 8


RESULT  11
AAW85522
ID   AAW85522 standard; peptide; 9 AA.
XX
AC   AAW85522;
XX
DT   22-APR 1998  (first entry)
XX
DE   Rat T-cell receptor peptide SEQ ID NO:40.
XX
KW   T cell stimulatory peptide; immunogen; non dendritic carrier; tumour;
KW   scaffold; inhibition; metastasis; wound healing; solid phase.
XX
OS   Rattus sp.
XX
PN   WO9748011-A1.
XX
ID   16-DEC 1997.
XX
PF   03-APR 1997;   97WO-DE00146.
XX
PR   03-APR 1996;   96DK-0000398.
XX
PA   (PEPR-) PEPRESEARCH AS.
XX
PI   Hoegaard PMH,  Jakobsen PH;
XX
DR   WPI; 1997-512645/47.
XX
PT   Non-dendritic peptide carrier linked to a solid phase useful as a
PT   diagnostic agent and as a scaffold for production of chemical
PT   derivatives
XX
PS   Example 27; Page 149; 262pp; English.
XX
CC   A non dendritic peptide carrier (A) has been developed which is coupled
CC   through a linker to a solid phase, forming a complex of (A) solid phase.
CC   Where (A) comprises 10-50 amino acids capable of forming a secondary
CC   structure in a benign buffer after liberation from the solid phase, and
CC   further the (A) solid phase complex comprises an immunogenic substance
CC   and/or an immune mediator coupled on (A). The present sequence
CC   represents a peptide used in an example from the present invention. An
CC   (A)-solid phase complex can be used as a scaffold for the production of
CC   chemical derivatives, characterised by covalently attaching molecules at
CC   attachment points. Alternatively (A) is used as a scaffold peptide for
CC   the incorporation into an Immunostimulating Complex (Iscom) resulting an
CC   (A)-Iscom complex which is used for the chemical coupling of antigenic
CC   substances in an aqueous solution by conjugation. (A) derivatised with
CC   one or more peptides having fibronectin-, laminin- or vitronectin-like
CC   binding activities can be used for the promotion of cell attachment to
CC   plastic surfaces, in particular to inhibit tumour growth and metastasis,
CC   and for promotion of wound healing. Also a derivatised (A) can be used
CC   for the selection of specifically-binding aptamers or as a diagnostic
CC   agent. Such diagnostic (A) molecules could be used to detect molecules
CC   derived from or indicative of pregnancy or of a disease such as an
CC   infectious, autoimmune or cancerous disease.
XX
SQ   Sequence   9 AA;


  Query Match             48.8%;  Score 21;  DB 18;  Length 9;
  Best Local Similarity   71.4%;  Pred. No. 6.4e+05;
  Matches    5;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       2 TLASSET 8
           | || ||
Db       2 TKASKET 8


RESULT  12
AAW85522
ID   AAW85522 standard; peptide; 9 AA.
XX
AC   AAW85522;
XX
DT   22-APR 1998  (first entry)
XX
DE   Rat T-cell receptor peptide SEQ ID NO:40.
XX
KW   T cell stimulatory peptide; immunogen; non dendritic carrier; tumour;
KW   scaffold; inhibition; metastasis; wound healing; solid phase.
XX
OS   Rattus sp.
XX
PN   WO9748011-A1.
XX
ID   16-DEC 1997.
XX
PF   03-APR 1997;   97WO-DE00146.
XX
PR   03-APR 1996;   96DK-0000398.
XX
PA   (PEPR-) PEPRESEARCH AS.
XX
PI   Hoegaard PMH,  Jakobsen PH;
XX
DR   WPI; 1997-512645/47.
XX
PT   Non-dendritic peptide carrier linked to a solid phase useful as a
PT   diagnostic agent and as a scaffold for production of chemical
PT   derivatives
XX
PS   Example 27; Page 149; 262pp; English.
XX
CC   A non dendritic peptide carrier (A) has been developed which is coupled
CC   through a linker to a solid phase, forming a complex of (A) solid phase.
CC   Where (A) comprises 10-50 amino acids capable of forming a secondary
CC   structure in a benign buffer after liberation from the solid phase, and
CC   further the (A) solid phase complex comprises an immunogenic substance
CC   and/or an immune mediator coupled on (A). The present sequence
CC   represents a peptide used in an example from the present invention. An
CC   (A)-solid phase complex can be used as a scaffold for the production of
CC   chemical derivatives, characterised by covalently attaching molecules at
CC   attachment points. Alternatively (A) is used as a scaffold peptide for
CC   the incorporation into an Immunostimulating Complex (Iscom) resulting an
CC   (A)-Iscom complex which is used for the chemical coupling of antigenic
CC   substances in an aqueous solution by conjugation. (A) derivatised with
CC   one or more peptides having fibronectin-, laminin- or vitronectin-like
CC   binding activities can be used for the promotion of cell attachment to
CC   plastic surfaces, in particular to inhibit tumour growth and metastasis,
CC   and for promotion of wound healing. Also a derivatised (A) can be used
CC   for the selection of specifically-binding aptamers or as a diagnostic
CC   agent. Such diagnostic (A) molecules could be used to detect molecules
CC   derived from or indicative of pregnancy or of a disease such as an
CC   infectious, autoimmune or cancerous disease.
XX
SQ   Sequence   9 AA;


  Query Match             48.8%;  Score 21;  DB 18;  Length 9;
```

Best Local Similarity 66.7%; Pred. No. 6.4e-05;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 4 ASSFEI 9
||:|:|
Db 1 ASSYSI 6

RESULT 13
AAW17496
ID AAW17496 standard; peptide; 9 AA.
XX
AC AAW17496;
XX
DT 01-APR-1997 (first entry)
XX
DE Protein kinase C-mu peptide mu V3-3 (457-465).
XX
KW Signal transduction; cell signalling; modulator; immune response;
KW protein kinase C; receptor for activated kinase C; RACK;
KW PKC mu; cognate; graft rejection; autoimmune disease;
KW allergy; asthma; therapy.
XX
OS Homo sapiens.
XX
PN WO9714436-A1.
XX
PD 17-APR-1997.
XX
PF 11-OCT-1996; 96WO-US16196.
XX
PR 12-OCT-1995; 95US-0542647.
PR 13-JUL-1995; 95US-0501964.
PR 31-JAN-1996; 96US-0594447.
XX
PA (JEFF-) JEFFERSON FOUNDATION INC.
XX
PI Kanner SM, Mochly-Rosen D, Napolitano EW, Ron D;
XX
DR WPI; 1997-236030/21.
XX
PT Identifying a modulator of intracellular signal transduction - by
PT determining the interaction of a signal generating peptide with the
PT test substance, allows modulation of the immune system
XX
PS Example 4; Page 43; 74pp; English.
XX
CC This sequence is a peptide, designated as V3-3, that corresponds to
CC amino acid residues 457-465 in the V3 region of protein kinase C
CC (PKC) mu. It was tested as a signal generating peptide in a
CC claimed method for identifying modulators of intracellular signal
CC transduction. This method assesses the ability of candidate
CC modulators to affect the interaction between a signal generating
CC protein, such as a PKC isozyme peptide (see AAW17722-7, AAW17741,
CC AAW15784-85, AAW17452-78), and a cognate binding protein involved in
CC modulating the signal transduction function. Identified substances
CC are useful as immunomodulators (claimed). They reduce T cell
CC activity, reduce the rate of graft rejection, reduce the severity
CC of an autoimmune disorder, ameliorate allergy and/or asthma, or
CC diminish a cytokine response (claimed).
XX
SQ Sequence 9 AA;

Query Match 50.0%; Score 21; DB 19; Length 9;
Best Local Similarity 50.0%; Pred. No. 6.4e-05;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LASSEI 7
|| :|:|
Db 2 LTSGDI 7

RESULT 14
AAY64540
ID AAY64540 standard; peptide; 9 AA.
XX
AC AAY64540;
XX
DT 02-MAR-2000 (first entry)
XX
DE Cadherin related neuronal receptor CAR cyclic peptide SEQ ID NO:38.
XX
KW Modulation; nonclassical cadherin mediated cell adhesion; CAR;
KW inhibition; cadherin extracellular domain; cell adhesion recognition;
KW OB cadherin; cadherin 5; cadherin 6; cadherin 7; cadherin 8;
KW cadherin 12; cadherin 14; cadherin 15; T cadherin; CR3 cadherin;
KW cadherin related neuronal receptor; FAT cadherin; protocadherin;
KW desmoglein; desmocollin; calcium binding; cancer; tumour; obesity;
KW rheumatoid arthritis; multiple sclerosis; diabetes; metastasis;
KW neurological disease; cyclic.
XX
OS Synthetic.
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Modified site 1..9
FT /note= "the termini residues are cyclised with each
FT other to form a cyclic peptide"
XX
PN WO9957149-A2.
XX
PD 11-NOV-1999.
XX
PF 05-MAY-1999; 99WO-CA000364.
XX
PR 05-MAY-1998; 98US-0083040.
PR 06-NOV-1998; 98US-0187859.
PR 20-JAN-1999; 99US-0234495.
PR 08-MAR-1999; 99US-0264516.
XX
PA (ADHE) ADHEREX TECHNOLOGIES INC.
XX
PI Blaschuk OW, Gour BJ, Byers S;
XX
DR WPI; 2000-038791/03.
XX
PT New cadherin modulating agents, used to modulate non-classical
PT cadherin mediated functions for treating e.g. cancers, obesity,
PT rheumatoid arthritis, multiple sclerosis, diabetes or neurological
PT disease.
XX
PS Claim 102; Page 220; 252pp; English.
XX
CC The present invention describes cadherin modulatory agents (MA)
CC comprising peptides which comprise a nonclassical cadherin cell adhesion
CC recognition (CAR) sequence. The MAs can be used to: modulate
CC nonclassical cadherin-mediated functions. They can be used for e.g.
CC inhibiting adhesion of nonclassical-cadherin expressing cells in a
CC mammal, enhancing delivery of a drug through the skin of a mammal,
CC enhancing delivery of a drug to a tumour in a mammal, treating cancer in
CC a mammal, inhibiting metastasis of a cancer in a mammal, inhibiting
CC angiogenesis in a mammal, inducing apoptosis in a non-classical cadherin
CC expressing cell, preventing or treating obesity in a mammal, stimulating
CC blood vessel regression in a mammal, enhancing drug delivery to the
CC central nervous system, treating a demyelinating neurological disease,
CC increasing vasopermeability in a mammal, enhancing adhesion of
CC nonclassical cadherin-expressing cells, inhibiting synaptic stability in
CC a mammal, or preventing pregnancy in a mammal. The MAs can also be used for
CC e.g. enhancing or directing neurite outgrowth, facilitating wound
CC healing or reducing scar tissue, or enhancing adhesion of tissue
CC in a mammal. They can also be used for treating e.g. psoriasis,
CC arthritis, age-related macular degeneration, multiple sclerosis and
CC diabetes. The products can also be used for detection and diagnosis and
CC in biosensors. AAY60502 to AAY64572 represent specifically claimed

```
;;   positions 1 to 9 of AAY64743 to AAY64944 and AA488283 to AA488483 represent
;;   sequences used in the exemplification of the present invention.
XX
SQ   Sequence   9 AA;


  Query Match             48.8%;  Score 21;  DB 22;  Length 9;
  Best Local Similarity   71.4%;  Pred. No. 5.4e+05;
  Matches    5;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 FTLASSE 7
             |:| ||
Db       1 FSLPSPE 7


Search completed: August 26, 2002, 06:27:08
Job time: 989 sec
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

Run on: August 26, 2002, 06:23:20 ; Search time 57.79 Seconds
(without alignments)
4.245 Million cell updates/sec

Title: US-09-744-282-2
Perfect score: 48
Sequence: 1 FTLASSETG 9

Scoring table: BLOSUM62
Xapop 10.0 , Xgapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: [illegible]

Minimum DB seq length: 0
Maximum DB seq length: 0

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/US8_COMB.pep *
2: /cgn2_6/ptodata/1/iaa/US9_COMB.pep *
3: /cgn2_6/ptodata/1/iaa/US10_COMB.pep *
4: /cgn2_6/ptodata/1/iaa/US11_COMB.pep *
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep *
6: /cgn2_6/ptodata/1/iaa/US60_COMB.pep *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	48	100.0	9	1	US-08-619-645-2	Sequence 2, Appli
2	48	100.0	9	2	US-08-644-493-2	Sequence 2, Appli
3	28	58.3	8	1	US-08-315-461-2	Sequence 2, Appli
4	28	58.3	9	1	US-08-619-645-1	Sequence 1, Appli
5	28	58.3	9	2	US-08-644-493-1	Sequence 1, Appli
6	21	43.8	9	3	US-09-187-859-1251	Sequence 1251, Ap
7	21	43.8	9	1	US-08-594-447-9	Sequence 9, Appl
8	21	43.8	9	1	US-08-541-964-4a	Sequence 4a, Appl
9	21	43.8	9	2	US-08-665-647-6	Sequence 6, Appl
10	21	43.8	9	4	US-09-187-859-3857	Sequence 3857, Ap
11	21	43.8	9	5	PCT US92-08094-8	Sequence 8, Appli
12	20	41.5	5	4	US-09-187-859-156	Sequence 156, App
13	20	41.5	5	4	US-09-187-859-1279	Sequence 1279, Ap
14	20	41.5	6	4	US-09-187-859-159	Sequence 159, App
15	20	41.5	6	4	US-09-187-859-1281	Sequence 1281, Ap
16	20	41.5	7	1	US-08-136-744-8-8	Sequence 8, Appl
17	20	41.5	7	4	US-09-040-216-20	Sequence 20, Appl
18	20	41.5	7	4	US-09-187-859-162	Sequence 162, App
19	20	41.5	7	4	US-09-187-859-1066	Sequence 1066, Ap
20	20	41.5	7	4	US-09-187-859-1113	Sequence 1113, Ap
21	20	41.5	7	4	US-09-187-859-1154	Sequence 1154, Ap
22	20	41.5	7	4	US-09-187-859-1196	Sequence 1196, Ap
23	20	41.5	7	4	US-09-187-859-1241	Sequence 1241, Ap
24	20	41.5	7	4	US-09-187-859-1284	Sequence 1284, Ap
25	20	41.5	8	4	US-09-187-859-165	Sequence 165, App
26	20	41.5	8	4	US-09-187-859-1069	Sequence 1069, Ap
27	20	41.5	8	4	US-09-187-859-1116	Sequence 1116, Ap
28	20	41.5	8	4	US-09-187-859-1157	Sequence 1157, Ap
29	20	41.5	8	4	US-09-187-859-1199	Sequence 1199, Ap
30	20	41.5	8	4	US-09-187-859-1244	Sequence 1244, Ap
31	20	41.5	8	4	US-09-187-859-1287	Sequence 1287, Ap
32	20	41.5	8	4	US-09-187-859-4748	Sequence 4748, Ap
33	20	41.5	9	1	US-08-584-224-2	Sequence 16, Appl
34	20	41.5	9	4	US-09-187-859-168	Sequence 168, App
35	20	41.5	9	4	US-09-187-859-1072	Sequence 1072, Ap
36	20	41.5	9	4	US-09-187-859-1119	Sequence 1119, Ap
37	20	41.5	9	4	US-09-187-859-1161	Sequence 1161, Ap
38	20	41.5	9	4	US-09-187-859-1202	Sequence 1202, Ap
39	20	41.5	9	4	US-09-187-859-1247	Sequence 1247, Ap
40	20	41.5	9	4	US-09-187-859-1290	Sequence 1290, Ap
41	20	41.5	9	4	US-09-187-859-3751	Sequence 3751, Ap
42	19	39.2	6	1	US-08-290-448A-41	Sequence 41, Appl
43	19	39.2	6	1	US-08-290-448A-41	Sequence 41, Appl
44	19	39.2	6	1	US-08-175-069A-41	Sequence 41, Appl
45	19	39.2	6	4	US-09-461-939A-41	Sequence 41, Appl

ALIGNMENTS

RESULT 1
US-08-619-645-2
; Sequence 2, Application US/08619645
; Patent No. 5736507
; GENERAL INFORMATION:
; APPLICANT: Boots, Anna M.H.
; APPLICANT: Verheijden, Gijsbertus F.M.
; TITLE OF INVENTION: No. 5736507o1 peptides derived from
; TITLE OF INVENTION: autoantigen for use in immunotherapy of autoimmune
; TITLE OF INVENTION: disease
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Appln No. 5736507o1 Patent Department
; STREET: 1300 Piccard Drive, Suite 206
; CITY: Rockville
; STATE: Maryland
; COUNTRY: U.S.A.
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/619,645
; FILING DATE: 25-MAR-1996
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Gormley, Mary E.
; REGISTRATION NUMBER: 34,409
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 947-4433
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-619-645-2

Query Match 100.0%; Score 48; DB 1; Length 9
Best Local Similarity 100.0%; Pred. No. [illegible]
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FTLASSETG 9
||||||||| |
Db 1 FTLASSETG 9

```
RESULT   2
US-08-315-461-2
; Sequence 2, Application US/08315461
; Patent No. 5861440
;  GENERAL INFORMATION:
;    APPLICANT:  A.M.H. Boots
;    APPLICANT:  G.F.M. Verheijden
;    APPLICANT:  F.H. Bos
;    TITLE OF INVENTION:  No. 5861440el Peptides derived from autoantigen for use
;    TITLE OF INVENTION:  in Immunotherapy of Autoimmune Diseases
;    NUMBER OF SEQUENCES:  10
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Akzo No. 5861440el Patent Department
;      STREET:  1300 Piccard Drive, Suite 206
;      CITY:  Rockville
;      STATE:  Maryland
;      COUNTRY:  USA
;      ZIP:  20850
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/315,461
;      FILING DATE:  Concurrently Herewith
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/215,645
;      FILING DATE:  25-MAR-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/EP94/01041
;      FILING DATE:  25-MAR-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  NL 9400531
;      FILING DATE:  31-MAR-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  NL 9400532
;      FILING DATE:  31-MAR-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Parker, David L.
;      REGISTRATION NUMBER:  32,165
;      REFERENCE/DOCKET NUMBER:  0M10.606
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (512) 418-3000
;      TELEFAX:  (512) 474-7577
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  8 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-315-461-2


  Query Match             54.5%;  Score 24;  DB 1;  Length 8;
  Best Local Similarity   71.4%;  Pred. No. 1.7e+05;
  Matches    5;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       2 LASSEI 8
             |:|||.|
Db       1 LVASSSI 7


RESULT   4
US-08-619-645-1
; Sequence 1, Application US/08619645
; Patent No. 5736507
;  GENERAL INFORMATION:
;    APPLICANT:  Boots, Anna M.H.
;    APPLICANT:  Verheijden, Gijsbertus F.M.
;    TITLE OF INVENTION:  No. 5736507el peptides derived from
;    TITLE OF INVENTION:  autoantigen for use in immunotherapy of autoimmune
;    TITLE OF INVENTION:  disease
;    NUMBER OF SEQUENCES:  9
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Akzo No. 5736507el Patent Department
;      STREET:  1300 Piccard Drive, Suite 206
;      CITY:  Rockville
;      STATE:  Maryland
;      COUNTRY:  U.S.A.
;      ZIP:  20850
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC DOS/MS DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25(EPO)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/619,645
;      FILING DATE:  25-MAR-1996
;      CLASSIFICATION:  514
```

```
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Dempsey, Mary E.
;       REGISTRATION NUMBER:  41,439
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (301) 947 4423
;   INFORMATION FOR SEQ ID NO:  1:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  9 amino acids
;       TYPE:  amino acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  peptide
US-08-744-282-1


  Query Match             54.5%;  Score 23;  DB 1;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 1.7e+05;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 FTLAS 5

Db       5 FTLAS 9


RESULT   5
US-08-834-449-1
; Sequence 1, Application US/08634493
; Patent No. 5843449
; GENERAL INFORMATION:
;     APPLICANT:  A.M.H. Boots
;     APPLICANT:  J.F.M. Verhijden
;     APPLICANT:  E.S. Bos
;     TITLE OF INVENTION:  No. 5843449e1 Peptides derived from and antibody for use
;     TITLE OF INVENTION:  in Immunotherapy of Autoimmune Diseases
;     NUMBER OF SEQUENCES:  10
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Akzo No. 5843449e1 Patent Department
;       STREET:  1300 Piccard Drive, Suite 206
;       CITY:  Rockville
;       STATE:  Maryland
;       COUNTRY:  USA
;       ZIP:  20850
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/834,493
;       FILING DATE:  18-APR-1996
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/015,445
;       FILING DATE:  25-MAR-1996
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  PCT/EP95/04201
;       FILING DATE:  25-OCT-1995
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  NL 9420l287
;       FILING DATE:  27-OCT-1994
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  NL 9520088O
;       FILING DATE:  07-APR-1995
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (301) 258-5200
;       TELEFAX:  (301) 977-0847
;   INFORMATION FOR SEQ ID NO:  1:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  9 amino acids
;       TYPE:  amino acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  peptide
US-08-834-449-1


  Query Match             54.5%;  Score 23;  DB 2;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 1.7e+05;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 FTLAS 5
           |||||
Db       5 FTLAS 9


RESULT   6
US-09-187-859-3854
; Sequence 3854, Application US/09187859A
; Patent No. 6358920
; GENERAL INFORMATION:
;  APPLICANT: Blaschuk, Orest W.
;  APPLICANT: Gour, Barbara J.
;  TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING NONCLASSICAL
;  TITLE OF INVENTION: CADHERIN MEDIATED FUNCTIONS
;  FILE REFERENCE: 100086.4070
;  CURRENT APPLICATION NUMBER: US/09/187,859A
;  CURRENT FILING DATE: 1998-11-06
;  NUMBER OF SEQ ID NOS: 4052
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3854
;   LENGTH: 8
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Representative cyclic modulating motif based on
;   OTHER INFORMATION: cadherin-related neuronal receptor cell adhesion
;   OTHER INFORMATION: recognition sequence
US-09-187-859-3854


  Query Match             48.8%;  Score 21;  DB 4;  Length 8;
  Best Local Similarity   57.1%;  Pred. No. 1.7e+05;
  Matches    4;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 FTLASSE 7
           ||: ||:
Db       2 FTIDSSD 8


RESULT   7
US-08-594-447-49
; Sequence 49, Application US/08594447
; Patent No. 5776716
; GENERAL INFORMATION:
;     APPLICANT:  Ron, Dorit
;     APPLICANT:  Napolitano, Eugene W.
;     APPLICANT:  Voronova, Anna F.
;     TITLE OF INVENTION:  METHODS FOR IDENTIFYING AGENTS WHICH
;     TITLE OF INVENTION:  MODULATE INTERACTION BETWEEN PKN-BETA-THETA, AND USES
;     TITLE OF INVENTION:  THEREOF
;     NUMBER OF SEQUENCES:  75
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  MORRISON & FOERSTER
;       STREET:  2000 Pennsylvania Avenue, NW - Ste. 5500
;       CITY:  Washington
;       STATE:  DC
;       COUNTRY:  USA
;       ZIP:  20006-1888
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.30
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/594,447
;       FILING DATE:  31-JAN-1996
;       CLASSIFICATION:  435
```

; ATTORNEY/AGENT INFORMATION:
; NAME: Murashige, Kate H.
; REGISTRATION NUMBER: 29,959
; REFERENCE/DOCKET NUMBER: 22550-20025.20
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 887-1500
; TELEFAX: (202) 822-0168
; TELEX: 90-4030 MRSNFOERSWSH
; INFORMATION FOR SEQ ID NO: 48:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 1..9
; OTHER INFORMATION: /label= mu V3 3
US-08-541-964-48

Query Match 48.8%; Score 21; DB 1; Length 9;
Best Local Similarity 50.0%; Pred. No. 1.7e+05;
Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 2 HASSEG 9
|| :|||
Db 2 HFQNDIG 9

RESULT 9
US-08-665-647-63
; Sequence 63, Application US/08665647
; Patent No. 5935804
; GENERAL INFORMATION:
; APPLICANT: Dasquez, Nicki J.
; APPLICANT: Ron, Dorit
; APPLICANT: Voronova, Anna F.
; APPLICANT: Napolitano, Eugene W.
; TITLE OF INVENTION: METHODS for IDENTIFY IMMUNOMODULATORS
; TITLE OF INVENTION: USING SIGNAL INTERACTION of PKC THETA
; NUMBER OF SEQUENCES: 89
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MCKINSON & FOERSTER
; STREET: 2000 Pennsylvania Avenue, NW Ste. 5500
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20006 1888
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC DOS/MS DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/665,647
; FILING DATE: 18 JUN 1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Murashige, Kate H.
; REGISTRATION NUMBER: 29,959
; REFERENCE/DOCKET NUMBER: 22550-20025.20
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 887-1500
; TELEFAX: (202) 822-0168
; TELEX: 90-4030 MRSNFOERSWSH
; INFORMATION FOR SEQ ID NO: 63:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 1..9
; OTHER INFORMATION: /label= mu V3 3
US-08-665-647-63

Query Match 48.8%; Score 21; DB 2; Length 9;
Best Local Similarity 50.0%; Pred. No. 1.7e+05;
Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 2 HASSEG 9
|| :|||

```
Db       2 TLEQNVIG 9


RESULT  10
US-09-187-859-3857
; Sequence 3857, Application US/09187859A
; Patent No. 6358920
; GENERAL INFORMATION:
;  APPLICANT: Blaschuk, Orest W.
;  APPLICANT: Gour, Barbara J.
;  TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING NONCLASSICAL
;  TITLE OF INVENTION: CADHERIN-MEDIATED FUNCTIONS
;  FILE REFERENCE: 100086.407C1
;  CURRENT APPLICATION NUMBER: US/09/187,859A
;  CURRENT FILING DATE: 1998-11-06
;  NUMBER OF SEQ ID NOS: 4052
;  SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 3857
;    LENGTH: 9
;    TYPE: PRT
;    ORGANISM: Artificial Sequence
;    FEATURE:
;     OTHER INFORMATION: Representative cyclic modulating agent based on
;     OTHER INFORMATION: cadherin related neuronal receptor cell adhesion
;     OTHER INFORMATION: recognition sequence
US-09-187-859-3857


  Query Match             48.8%;  Score 21;  DB 4;  Length 9;
  Best Local Similarity   57.1%;  Pred. No. 1.7e-05;
  Matches    4;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 FIASSE 7
           ||:|:|
Db       3 FIDSSD 9


RESULT  11
PCT-US92-08094-3
; Sequence 3, Application PC/TUS9208094
;  GENERAL INFORMATION:
;     APPLICANT:  GENENTECH, INC.
;     APPLICANT:  Amento, Edward P.
;     TITLE OF INVENTION:  DIAGNOSING AND TREATING ATHEROSCLEROSIS
;     TITLE OF INVENTION:  DISORDERS
;     NUMBER OF SEQUENCES:  80
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Genentech, Inc.
;       STREET:  460 Point San Bruno Blvd
;       CITY:  South San Francisco
;       STATE:  California
;       COUNTRY:  USA
;       ZIP:  94 080-4990
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  5.25 inch, 360 Kb floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC DOS/MS-DOS
;       SOFTWARE:  patin (Genentech)
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  PCT/US92/08094
;       FILING DATE:  19920923
;       CLASSIFICATION:
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  07/765222
;       FILING DATE:  23-SEP-1991
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  07/773145
;       FILING DATE:  18-OCT-1991
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  07/854362
;       FILING DATE:  18-MAR-1992
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Hensley, Max D.
;       REGISTRATION NUMBER:  27,043
;       REFERENCE/DOCKET NUMBER:  734P3
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  415/225-1994
;       TELEFAX:  415/952-9881
;       TELEX:  910/371-7168
;  INFORMATION FOR SEQ ID NO:  3:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  9 amino acids
;       TYPE:  AMINO ACID
;       TOPOLOGY:  linear
PCT-US92-08094-3


  Query Match             48.8%;  Score 21;  DB 5;  Length 9;
  Best Local Similarity   66.7%;  Pred. No. 1.7e-05;
  Matches    4;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       4 ASSETG 9
           |||::|
Db       1 ASSDSG 6


RESULT  12
US-09-187-859-156
; Sequence 156, Application US/09187859A
; Patent No. 6358920
; GENERAL INFORMATION:
;  APPLICANT: Blaschuk, Orest W.
;  APPLICANT: Gour, Barbara J.
;  TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING NONCLASSICAL
;  TITLE OF INVENTION: CADHERIN MEDIATED FUNCTIONS
;  FILE REFERENCE: 100086.407C1
;  CURRENT APPLICATION NUMBER: US/09/187,859A
;  CURRENT FILING DATE: 1998-11-06
;  NUMBER OF SEQ ID NOS: 4052
;  SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 156
;    LENGTH: 5
;    TYPE: PRT
;    ORGANISM: Artificial Sequence
;    FEATURE:
;    OTHER INFORMATION: Representative linear modulating agent based on
;    OTHER INFORMATION: cadherin-6 cell adhesion recognition sequence
US-09-187-859-156


  Query Match             46.5%;  Score 20;  DB 4;  Length 5;
  Best Local Similarity   100.0%;  Pred. No. 1.7e-05;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       6 SETG 9
           ||||
Db       2 SEIG 5


RESULT  13
US-09-187-859-1279
; Sequence 1279, Application US/09187859A
; Patent No. 6358920
; GENERAL INFORMATION:
;  APPLICANT: Blaschuk, Orest W.
;  APPLICANT: Gour, Barbara J.
;  TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING NONCLASSICAL
;  TITLE OF INVENTION: CADHERIN-MEDIATED FUNCTIONS
;  FILE REFERENCE: 100086.407C1
;  CURRENT APPLICATION NUMBER: US/09/187,859A
;  CURRENT FILING DATE: 1998-11-06
;  NUMBER OF SEQ ID NOS: 4052
;  SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 1279
;    LENGTH: 5
;    TYPE: PRT
```

; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Representative cyclic modulating agent based on
; OTHER INFORMATION: cadherin 6 cell adhesion recognition sequence

Query Match 46.5%; Score 20; DB 4; Length 6;
Best Local Similarity 100.0%; Pred. No. 1.7e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 SEIG 9
||||
Db 6 SEIG 6

RESULT 2
US-09-744-282-2
; Sequence 2, Application US/09187859A
; Patent No. 6459
; GENERAL INFORMATION:
; APPLICANT: Blaschuk, Orest W.
; APPLICANT: Doyle, Barbara J.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING N-CLASSICAL
; TITLE OF INVENTION: CADHERIN MEDIATED FUNCTIONS
; FILE REFERENCE: 10000B6.40/C1
; CURRENT APPLICATION NUMBER: US/09/187,859A
; CURRENT FILING DATE: 1998-11-06
; NUMBER OF SEQ ID NOS: 4052
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 2
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Representative linear modulating agent based on
; OTHER INFORMATION: cadherin 6 cell adhesion recognition sequence

Query Match 46.5%; Score 20; DB 4; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.7e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 SEIG 9
||||
Db 6 SEIG 9

RESULT 3
US-09-744-282-2
; Sequence 1281, Application US/09187859A
; Patent No. 6459
; GENERAL INFORMATION:
; APPLICANT: Blaschuk, Orest W.
; APPLICANT: Doyle, Barbara J.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING N-CLASSICAL
; TITLE OF INVENTION: CADHERIN MEDIATED FUNCTIONS
; FILE REFERENCE: 10000B6.40/C1
; CURRENT APPLICATION NUMBER: US/09/187,859A
; CURRENT FILING DATE: 1998-11-06
; NUMBER OF SEQ ID NOS: 4052
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 1281
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Representative cyclic modulating agent based on
; OTHER INFORMATION: cadherin 6 cell adhesion recognition sequence

Query Match 46.5%; Score 20; DB 4; Length 6;
Best Local Similarity 100.0%; Pred. No. 1.7e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 SEIG 9
||||
Db 6 SEIG 6

Search completed: August 26, 2002, 06:23:21
Job time: 762 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

Algorithm: protein search, using sw model

Run on: August 26, 2002, 06:24:48 ; Search time 7.+ Seconds
(without alignments)
11,702 Million cell updates/sec

Title: US-09-744-282-3
Perfect score: 46
Sequence: 1 YDDQESVKS 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283108 seqs, 96086334 residues

Total number of hits satisfying chosen parameters: 566216

Minimum DB seq length: 0
Maximum DB seq length: 9

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_71:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	20	43.5	7	2	B39040	calsequestrin, fas
2	19	41.3	9	2	PW0002	chlo; gb:; a/b-bi
3	18	39.1	7	2	A34818	vinculin 72K chain
4	15	32.6	7	2	S68004	[acyl-in 75K chain
5	14	30.4	5	2	A32014	transmembrane Esc
6	14	30.4	8	2	A61428	trypsin (EC 3.4.21
7	14	30.4	9	2	C24180	tibetanosen beta ch
8	14	30.4	9	2	PH0108	late 51 kD protein
9	13	28.3	4	2	A41890	protein - Escher
10	13	28.3	5	2	PQ0089	phosphatase 1 (EC 3
11	13	28.3	5	2	PH0308	50 kDa heavy chain CH0
12	13	28.3	6	2	PH0503	c-erbB receptor to
13	13	28.3	7	2	S07234	[acyl-carrier] p
14	13	28.3	7	2	I46868	alpha myosin heavy
15	13	28.3	7	2	PC0526	[acyl-carrier] p
16	13	28.3	8	2	S11078	[acyl-carrier] phosphat
17	13	28.3	8	2	PH0030	inulinase (EC 3.2.
18	13	28.3	8	2	B47303	beta peptide chain
19	13	28.3	8	2	S29272	transcription binding
20	13	28.3	9	2	PH0270	50 kDa heavy chain CH0
21	13	28.3	9	2	PH0272	50 kDa heavy chain CH0
22	13	28.3	9	2	S18784	chorion protein - bovin
23	13	28.3	9	2	S36850	50 kDa heavy chain V i
24	12	26.1	4	2	I40097	tissue A - chicken
25	12	26.1	5	2	PH0601	[acyl-carrier] p
26	12	26.1	5	2	PH0679	[acyl-carrier] p
27	12	26.1	6	2	A31263	[acyl-carrier] phosphat
28	12	26.1	6	2	B31263	dihydrofolate redu
29	12	26.1	6	2	A19780	transferrin - bovi
30	12	26.1	7	2	PH0628	[acyl-carrier] p
31	12	26.1	7	2	PH0722	[acyl-carrier] p
32	12	26.1	8	2	PC4130	hypothetical prote
33	12	26.1	8	2	PH0411	50 kDa heavy chain CH0
34	12	26.1	8	2	PH0507	c-erbB receptor pr
35	12	26.1	8	2	B47904	3-phosphoinositide-
36	12	26.1	8	2	A61499	3-phosphoinositide-
37	12	26.1	9	2	D28854	[acyl-carrier-protein] p
38	12	26.1	9	2	E28854	[acyl-carrier-protein] p
39	12	26.1	9	2	F28854	[acyl-carrier-protein] p
40	12	26.1	9	2	JN0029	sperm-specific cation c
41	12	26.1	9	2	JN0027	[acyl-carrier-protein] p
42	12	26.1	9	2	B47972	[acyl-carrier-protein] p
43	12	26.1	9	2	I46026	58 kDa keratin, t
44	12	26.1	9	2	PH0445	3-oxoacyl-[acyl-carrier
45	12	26.1	9	2	PH0502	c-erbB receptor pr

ALIGNMENTS

RESULT 1
B39040
calsequestrin, fast skeletal muscle - rabbit (fragment
C;Species: oryctolagus cuniculus (domestic rabbit)
C;Date: 23-Aug-1991 #sequence_revision 23-Aug-1991 #text_change 15-Feb-1995
C;Accession: B39040
R;Cala, S.E.; Jones, L.R.
J. Biol. Chem. 266, 391-398, 1991
A;Title: Phosphorylation of cardiac and skeletal muscle calsequestrin by the
A;Reference number: A39040; MUID:91016183
A;Accession: B39040
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-7 <CAL>
C;Keywords: phosphoprotein; skeletal muscle

Query Match 43.5%; Score 20; DB 2; Length 7;
Best Local Similarity 50.0%; Pred. No. 2.9e-05;
Matches 4; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 2 DDQESVK 7
| | | | | | | |
Db 1 DGEEDLK 6

RESULT 2
PW0002
chlorophyll a/b binding protein 21K - green algae (Dunaliella tertiolecta) (fragment)
N;Alternate names: photosystem II light harvesting chlorophyll 21K protein
C;Species: Dunaliella tertiolecta
C;Date: 24-Nov-1999 #sequence_revision 24-Nov-1999 #text_change 24-Nov-1999
C;Accession: PW0002
R;LaRoche, J.; Bennett, J.; Falkowski, P.G.
Gene 95, 165-171, 1990
A;Title: Characterization of a cDNA encoding for the 20.5 kDa LHCII apoprotein in m
A;Reference number: JW0040; MUID:91067628
A;Accession: PW0002
A;Molecule type: protein
A;Residues: 1-9 <LAR>
C;Superfamily: chlorophyll a/b binding protein
C;Keywords: chloroplast; grana; light-harvesting complex; membrane; photosynt

Query Match 41.3%; Score 19; DB 2; Length 9;
Best Local Similarity 50.0%; Pred. No. 1.2e-04;
Matches 2; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Qy 1 YDDQESV 7
| | | | | | | |
Db 3 YDEMQAM 9

RESULT 5
A43914
[illegible] K chain - pigeon pea (fragment)
[illegible] Cajanus cajan (pigeon pea)
C;Date: 15-Dec-1990 #sequence_revision 13-Jul-1990 #text_change 09-Sep-1994
C;Accession: A43914
R;Miwa, Y.K.; Mawal, M.R.; Kanyokar, P.K.
Biochem. Biophys. Res. Commun. 166, 1446-1451, 1990
A;Title: Thermal denaturation properties of viroid from Cajanus cajan.
A;Reference number: A43914; MUID:90167965
A;Accession: A43914
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-7 <MAW>

Query Match 41.8%; Score 16; DB 2; Length 7;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 DQQE 5
:::: |
Db 2 NDEE 5

RESULT 6
A61328
trypsin (EC 3.4.21.4) precursor - minke whale (fragment)
C;Species: Balaenoptera acutorostrata (minke whale, lesser rorqual)
C;Date: 17-Jul-1994 #sequence_revision 17-Jul-1994 #text_change 09-May-1999
C;Accession: A61328
R;Brieteux-Gregoire, S.; Schyns, R.; Florkin, M.; Dhanens, M.; Welling, G.W.; Beintema
Biochim. Biophys. Acta 486, 244-255, 1977
A;Title: N-terminal amino acid sequence of trypsinogen from the lesser rorqual, Balae
ceras.
A;Reference number: A61328; MUID:75146765
A;Accession: A61328
A;Molecule type: protein
A;Residues: 1-8 <BRI>
C;Keywords: hydrolase; protein digestion; serine proteinase; zymogen
F;1-8/Domain: activation peptide #status experimental <APP>

Query Match 30.4%; Score 14; DB 2; Length 8;
Best Local Similarity 50.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 DQQE 5
:: :
Db 4 DSSD 7

RESULT 7
C24180
fibrinogen beta chain - Japanese macaque (fragment)
N;Contains: fibrinopeptide B
C;Species: Macaca fuscata (Japanese macaque)
C;Date: 05-Jun-1988 #sequence_revision 10-Mar-1994 #text_change 26-Jan-1996
C;Accession: C24180
R;Nakamura, S.; Takenaka, O.; Takahashi, K.
J. Biochem. 97, 1487-1492, 1985
A;Title: Fibrinopeptides A and B of Japanese monkey (Macaca fuscata) and patas monkey
aemons, and baboons.
A;Reference number: A91990; MUID:85280140
A;Accession: C24180
A;Molecule type: protein
A;Residues: 1-9 <NAK>
C;Superfamily: fibrinogen beta chain; fibrinogen beta/gamma homology; fibrinogen disu

Query Match 30.4%; Score 14; DB 2; Length 9;
Best Local Similarity 42.9%; Pred. No. 2.8e+05;
Matches 3; Conservative 3; Mismatches 1; Indels 2; Gaps 2;

Qy 3 DQESVKS 9
::|| : |
Db 1 NEESLFS 7

RESULT 8
PH0108
late G1-69 protein - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 16-Jul-1999
C;Accession: PH0108
R;Nikaido, I.; Bradley, D.W.; Bardee, A.B.
Exp. Cell Res. 192, 102-109, 1991
A;Title: Molecular cloning of transcripts that accumulate during the late G1 phase in
A;Reference number: PH0108; MUID:91078351
A;Accession: PH0108
A;Status: preliminary
A;Molecule type: mRNA

A;Residues: 1 <NEK>

Query Match 30.4%; Score 14; DB 2; Length 9;
Best Local Similarity 66.7%; Pred. No. 2.8e+05;
Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 EQE 4
:||:
Db 8 DQE 6

RESULT 9
A41890
protein I - Escherichia coli (fragment)
C;Species: Escherichia coli
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 30-Sep-1993
C;Accession: A41890
R;Slattery, A.; Sebbardt, K.; Kristiansen, E.; Birkeland, N.K.; Lindqvist, B.H.
J. Bacteriol. 174, 4094-4100, 1992
A;Title: Escherichia coli K 12 and B contain functional beta-glucuronide P.-uptake genes.
A;Reference number: A41890; MUID:92283767
A;Accession: A41890
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-4 <SLA>
A;Cross-references: EMBL:M91463

Query Match 28.3%; Score 13; DB 2; Length 4;
Best Local Similarity 50.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 EQE 5
:||:
Db 1 EDE 4

RESULT 10
PQ0189
photosystem I 19.4K PI chain - common tobacco (fragment)
C;Species: Nicotiana tabacum (common tobacco)
C;Date: 30-May-1994 #sequence_revision 19-May-1994 #text_change 17-Mar-1995
C;Accession: PQ0189
R;Kobayashi, T.; Mekawa, K.; Hayashida, N.; Nakamura, M.; Shinozaki, K.
Plant Physiol. 102, 1259-1267, 1993
A;Title: Molecular heterogeneity of photosystem I. psaD, psaE, psaH and psaI are a
A;Reference number: PQ0067; MUID:94105345
A;Accession: PQ0189
A;Molecule type: protein
A;Residues: 1-55 <KOB>
C;Keywords: chloroplast; photosynthesis; photosystem I; thylakoid

Query Match 28.3%; Score 13; DB 2; Length 5;
Best Local Similarity 50.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YEQ 3
:||
Db 2 YDK 5

RESULT 11
PIB588
Ig heavy chain CDR3 region (clone 6-88) - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 30-Sep-1995 #sequence_revision 30-Sep-1995 #text_change 16-Aug-1996
C;Accession: PIB588
R;Yamada, M.; Wasserman, R.; Reichard, B.A.; Shane, S.; Caton, A.J.; Rovera, G.
J. Exp. Med. 173, 395-407, 1991
A;Title: Preferential utilization of specific immunoglobulin heavy chain diversity and
A;Reference number: PI0222; MUID:91108337
A;Accession: PI0308
A;Molecule type: DNA
A;Residues: 1-5 <YAM>
A;Experimental source: B lymphocyte
C;Keywords: heterotetramer; immunoglobulin

Query Match 28.3%; Score 13; DB 2; Length 5;
Best Local Similarity 75.0%; Pred. No. 2.8e+05;
Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 3 DQES 6
:|:|
Db 1 DWES 4

RESULT 12
PI0533
T cell receptor beta chain V D J region (126-2AA) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 31-May-1996
C;Accession: PI0533
R;Feeney, A.J.
J. Exp. Med. 174, 1115-124, 1991
A;Title: Junctional sequences of fetal T cell receptor beta chains live how N-term, as
A;Reference number: PI0509; MUID:91277?-1
A;Accession: PI0533
A;Status: translation not shown
A;Molecule type: mRNA
A;Residues: 1-6 <FEE>
A;Experimental source: adult thymus, strain BALB/c
C;Keywords: T cell receptor

Query Match 28.3%; Score 13; DB 2; Length 6;
Best Local Similarity 66.7%; Pred. No. 2.8e+05;
Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 DDQ 4
:||:
Db 4 DDR 6

RESULT 13
S57274
triacylglycerol lipase (EC 3.1.1.3) - Psychrobacter immobilis (fragment)
C;Species: Psychrobacter immobilis
C;Date: 28-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 03-Nov-1995
C;Accession: S57274
R;Arpigny, J.L.; Feller, G.; Gerday, C.
Biochim. Biophys. Acta 1264, 161, 1995
A;Title: Corrigendum to "Cloning, sequence and structural features of a lipase fr m t
A;Reference number: S57274; MUID:95459197
A;Accession: S57274
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-7 <ARP>
A;Cross-references: EMBL:X62712
C;Keywords: carboxylic ester hydrolase

Query Match 28.3%; Score 13; DB 2; Length
Best Local Similarity 14.3%; Pred. No. 2.8e+05;
Matches 1; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YDDQESV 7
: |::.:
Db 1 WGDKDQI 7

RESULT 14
146868
alpha-myosin heavy chain - rabbit (fragment)

Species: Oryctolagus cuniculus (domestic rabbit)
Date: 11-Feb-1997 #Sequence_revision 14-Feb-1997 #text_change 05-Nov-1999
Accession: A: 149868
Reference: 1; A; Sheela, P.K.; Sinha, A.M.; Basu, H.
Title: Nucl. Acid. Res. 0:3:A:81, 1984 9:44 9:48, 1984
Article: Characterization of genomic clones specifying rabbit alpha and beta ventricula
Reference Number: 146868; MEDLINE:84221901
Accession: A: 149868
Status: preliminary; translated from GB/EMBL/DDBJ
Molecule type: mRNA
Residues: 1-*
Cross-references: GB:K01698; NID:g155538; PIDN:AAA41415.1; PID:g155538
Query Match 28.4%; Score 14; DB 2; Length 2;
Best Local Similarity 66.7%; Pred. No. 2.7e+05;
Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
Qy 6 CgE 5
Db 1 CFE 3
RESULT 2
P10479
T-cell receptor beta chain V-D-J region (141.3) - mouse (fragment)
Species: Mus musculus (house mouse)
Date: 21-Jul-1989 #Sequence_revision 17-Jul-2002 #text_change 30-May-2002
Accession: P10479
Reference: A; 1
Eur. Exp. Med. Immunol. 174, 125-124, 1991
Title: Sequences of total T-cell receptor beta-chain gene V-N-regions.
Accession: P10479; MEDLINE:91277041
Accession: P10479
Status: translation not shown.
Molecule type: mRNA
Residues: 1-*
Experimental Source: day 19 fetal thymus; strain BALB/c
Keywords: T-cell receptor
Query Match 28.4%; Score 14; DB 2; Length 2;
Best Local Similarity 66.7%; Pred. No. 2.7e+05;
Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
Qy 6 CgE 4
1
Db 1 CFE 3

Search completed: August 26, 2002, 06:24:49
Job time: 5.1 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

SW protein: protein search, using sw model

Run on: August 26, 2002, 06:40:43 ; Search time 11.04 Seconds
(without alignments)
7,754 Million cell updates/sec

Title: US-09-744-282-3
Perfect score: 19
Sequence: 1 YDQESVKS 9

Scoring table: BL SUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0
Maximum DB seq length: 9

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

Prod. N. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	23	121.6	8	1	UPA1_HUMAN	P30087 homo sapien
2	16	84.2	9	1	NSK1_SARBU	P41492 sarcophaga
3	13	68.4	5	1	IRM3_ECOLI	P11673 escherichia
4	14	73.4	6	1	ASP2_LACSN	P82650 lactobacill
5	14	73.4	9	1	FIBB_MACFA	P2 47 macaca fasc
6	13	68.4	7	1	ALL7_CYDPO	P82158 cydia pomon
7	13	68.4	8	1	ALL5_CALVO	P41641 calliphora
8	13	68.4	8	1	ALL5_CYDPO	P82157 cydia pomon
9	12	63.1	5	1	BIPA_CITFR	P13071 citrobacter
10	12	63.1	9	1	FIBB_PAPAN	P21 34 papio anubi
11	12	63.1	9	1	FIBB_PAPHA	P13 16 papio hamad
12	12	63.1	9	1	FIBB_THEGE	P19 54 theropithec
13	12	63.1	9	1	MOSF_CLYJA	P19 57 clyporaster
14	12	63.1	9	1	MOSH_CLYJA	P19 52 clyporaster
15	12	63.1	9	1	UHA2_HUMAN	P47 39 homo sapien
16	11	57.9	7	1	WWA2_ACHFU	P47 23 achatina fu
17	11	57.9	8	1	LCK3_LEUMA	P21112 leucophaea
18	11	57.9	8	1	LCK6_LEUMA	P21113 leucophaea
19	11	57.9	8	1	UC26_MAIZE	P09 2 zea mays (m
20	11	57.9	8	1	UHO9_RAT	P56575 rattus norv
21	11	57.9	9	1	LPCA_STAAU	P46 54 staphyloco
22	1	21.7	7	1	EIO5_LITRU	P47101 litoria rub
23	1	21.7	7	1	UC24_MAIZE	P09 2 zea mays (m
24	1	21.7	7	1	UFO3_MOUSE	P09 41 mus musculu
25	1	21.7	8	1	CLP_THICH	P80438 thiobacillu
26	1	21.7	8	1	RSI_ERWCH	P47 95 erwinia cha
27	1	21.7	9	1	DCML_PSEPU	P19 61 pseudomonas
28	1	21.7	9	1	PLA2_TRIBY	P80 59 trimeresu b
29	1	21.7	9	1	PISI_BOMMO	P14 19 bombyx mori
30	1	21.7	9	1	UF02_MOUSE	P49 10 mus musculu
31	0	16.6	7	1	FAR3_HAECO	P52 08 haemonchus
32	0	16.6	7	1	FAR3_PANRE	P14 74 panagrellus
33	0	16.6	7	1	IGA2_DACDE	P49 04 dactylosom d
34	0	16.6	8	1	NSK1_MYCTU	P48 15 mycobacter
35	0	16.6	8	1	ORMY_ORCCI	P82 48 oreomertes
36	0	16.6	8	1	PLP_BRANA	P81 67 brassica na
37	0	16.6	9	1	FAR4_CALVI	P43 09 calliphora
38	0	16.6	9	1	FIBB_ERYPA	P29 49 erythrocebu
39	0	16.6	9	1	NEF_HV1Z8	P 1 61 human immu
40	0	16.6	9	1	SAMF_MUSCA	P29 95 musca dom a
41	0	16.6	9	1	TBYE_PIG	P01 55 sus scrofa
42	0	16.6	9	1	VLAI_HUMAN	P01 1 homo sapien
43	0	16.6	9	1	VLAH_HUMAN	P19 84 homo sapien
44	0	16.6	9	1	UN19_CLOPA	P01 55 cl striadum
45	0	16.6	9	1	UXE_VIBCI	P42 72 vibri 1 27

ALIGNMENTS

RESULT 1
UPA1_HUMAN
ID UPA1_HUMAN STANDARD; PRT; 9 AA.
AC P30087;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Unknown protein from 2D-page of plasma (spot 1) (fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE.
RC TISSUE=Plasma;
RX MEDLINE=93092947; PubMed=1456097;
RA Hughes G.J., Frutiger S., Paquet N., Bairoch F., Lasqu J.C.,
RA Sanchez J.-C., James R., Tissot J.-D., Hochstrasser D.F.;
RT "Plasma protein map: an update by microsequencing.";
RL Electrophoresis 13:707-714(1992).
CC -!- MISCELLANEOUS: ON THE 2-DEL GEL THE DETERMINED PI VALUE WAS UNKNOWN
CC PROTEIN IS: 4.9, ITS MW IS: 95 kDa.
DR SWISS-2DPAGE; P30087, HUMAN.
FT NON_TER 1 1
FT UNSURE 8 8
FT NON_TER 8 8
SQ SEQUENCE 8 AA; 964 MW; C5177C5F5B3C7CA CRC64;

Query Match 63.5%; Score 20; DB 1; Length 8;
Best Local Similarity 100.0%; Pred. No. 2.6e-3;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 DQES 6
||||
Db 1 DQES 4

RESULT 2
NSK1_SARBU
ID NSK1_SARBU STANDARD; PRT; 9 AA.
AC P41492;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-FEB-1996 (Rel. 33, Last annotation update)
DE Neosullakinin I (NEP-SK 1).
OS Sarcophaga bullata (Grey flesh fly) (Neobellieria bullata).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Oestroidea; Sarcophagidae; Sarcophaga.
OX NCBI_TaxID=7385;
RN [1]
RP SEQUENCE.
RC TISSUE=Head;
RX MEDLINE=94083161; PubMed=1906472;

```
RA   [illegible] A., Schoofs L., Proost P., Van Damme J., Section 2 A.;
RT   "Location and primary structure of two sulfakinin-like peptides from
RT   the locust, Neodelineria bellata.";
RL   Comp. Biochem. Physiol. 103C:139-142(1992).
CC   -!- FUNCTION: MYOTROPIC PEPTIDE.
CC   -!- SIMILARITY: BELONGS TO THE GASTRIN/CCK/LEVITICHIN FAMILY.
DR   InterPro; IPR001651; Gastrin.
DR   PROSITE; PS00264; GASTRIN; 1.
KW   Neuropeptide; Amidation; Sulfation.
FT   PEPTIDE       3       4       SULFATION (POTENTIAL).
FT   PEPTIDE       7       9       AMIDATION (POTENTIAL).
SQ   SEQUENCE   9 AA;  1107 MW;  [illegible] CRC64;


  Query Match             57.6%;  Score 15;  DB 1;  Length 9;
  Best Local Similarity   66.7%;  Pred. No. 1e+05;
  Matches    2;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        2 YDD 4
            |:|
Db        4 YNG 6


RESULT 5
IBAM_ECOLI
ID   IBAM_ECOLI     STANDARD;      PRT;     5 AA.
AC   P[illegible];
DT   01-JAN-1990 (Rel. 18, Created)
DT   01-JAN-1990 (Rel. 18, Last sequence update)
DT   15-JUL-2000 (Rel. 37, Last annotation update)
DE   IbaM protein (Fragment).
GN   IBAM
OS   Escherichia coli.
OC   Plasmid IncFII R1nf.
OC   Bacteria; Proteobacteria; Gamma subdivision; Enterobacteriaceae;
OC   Escherichia.
OX   NCBI_TaxID=562;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=90225[illegible]; PubMed=2183659;
RA   Inouye S., Yoshioka Y., Ohtsubo E.;
RT   "Identification and characterization of the products from the IraI
RT   and IraD genes of plasmid R100.";
RL   J. Bacteriol. 170:2749-2752(1988).
CC   -!- FUNCTION: NO TRANSFER GENE PRODUCT, IS INVOLVED IN THE EXCLUSION
CC       PROCESS OF BACTERIAL CELLS FOR THE EXCHANGE OF PLASMID DNA.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC   -!- SIMILARITY: TO IRAM PROTEIN OF OTHER PLASMIDS.
CC   -----------------------------------------------------------------------
CC   Swiss-Prot entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://[illegible]
CC   [illegible] for details).
CC   -----------------------------------------------------------------------
DR   EMBL; M[illegible]; -; NOT_ANNOTATED_CDS.
DR   PIR; A[illegible] A[illegible]4.
KW   Conjugation; Plasmid; DNA binding.
FT   NON_TER       1       1
SQ   SEQUENCE   5 AA;  634 MW;  [illegible] CRC64;


  Query Match             30.4%;  Score 14;  DB 1;  Length 5;
  Best Local Similarity   50.0%;  Pred. No. 1e+05;
  Matches    2;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        2 YDD 4
            |:|
Db        3 NEE 5


RESULT 4
ASP2_LACSN
ID   ASP2_LACSN     STANDARD;      PRT;     6 AA.
AC   P82655;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Acid shock protein 2 (Fragment).
OS   Lactobacillus sanfranciscensis (Lactobacillus sanfrancisco).
OC   Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillaceae;
OC   Lactobacillus.
OX   NCBI_TaxID=1625;
RN   [1]
RP   SEQUENCE.
RC   STRAIN=CB1;
RX   MEDLINE=21322712; PubMed=11429463;
RA   De Angelis M., Bini L., Pallini V., Cocconcelli P.S., Gobbetti M.;
RT   "The acid stress response in Lactobacillus sanfranciscensis CB1.";
RL   Microbiology 147:1863-1873(2001).
CC   -!- INDUCTION: OVEREXPRESSED IN ACID ENVIRONMENTS.
FT   NON_TER       6       6
SQ   SEQUENCE   6 AA;  778 MW;  6AA45B5B142A8000 CRC64;


  Query Match             30.4%;  Score 14;  DB 1;  Length 6;
  Best Local Similarity   66.7%;  Pred. No. 1e+05;
  Matches    2;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        1 YDD 3
            |:|
Db        4 YNG 6


RESULT 6
FIBB_MACFU
ID   FIBB_MACFU     STANDARD;      PRT;     9 AA.
AC   P19845;
DT   01-NOV-1990 (Rel. 16, Created)
DT   01-NOV-1990 (Rel. 16, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
DE   Fibrinogen beta chain [Contains: Fibrinopeptide B; (Fragment).
GN   FGB.
OS   Macaca fuscata fuscata (Japanese macaque).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
OC   Cercopithecinae; Macaca.
OX   NCBI_TaxID=9543;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=85289140; PubMed=3928610;
RA   Nakamura S., Takenaka O., Takahashi K.;
RT   "Fibrinopeptides A and B of Japanese monkey (Macaca fuscata) and
RT   patas monkey (Erythrocebus patas): their amino acid sequences,
RT   restricted mutations, and a molecular phylogeny for macaques,
RT   guenons, and baboons.";
RL   J. Biochem. 97:1487-1492(1985).
CC   -!- FUNCTION: FIBRINOGEN HAS A DOUBLE FUNCTION: YIELDING MONOMERS THAT
CC       POLYMERIZE INTO FIBRIN AND ACTING AS A CLEARER IN PLATELET
CC       AGGREGATION.
CC   -!- SUBUNIT: HEXAMER CONTAINING 2 SETS OF 3 NON-IDENTICAL CHAINS
CC       (ALPHA, BETA AND GAMMA), LINKED BY DISULFIDE BONDS.
CC   -!- MISCELLANEOUS: CONVERSION OF FIBRINOGEN TO FIBRIN IS TRIGGERED BY
CC       THROMBIN, WHICH CLEAVES FIBRINOPEPTIDES A AND B FROM ALPHA & BETA
CC       CHAINS, AND THUS EXPOSES THE N-TERMINAL POLYMERIZATION SITES
CC       RESPONSIBLE FOR THE FORMATION OF THE SOFT CLOT.
DR   PIR; C24180; C24180.
DR   InterPro; IPR002181; Fibrinogen_C.
DR   PROSITE; PS00514; FIBRIN_BC_C_DOMAIN; PARTIAL.
KW   Blood coagulation; Plasma.
FT   PEPTIDE       1       9       FIBRINOPEPTIDE B.
FT   NON_TER       9       9
SQ   SEQUENCE   9 AA;  1038 MW;  69FE65B9C735BB1B CRC64;
```

```
  Query Match             30.4%;  Score 14;  DB 1;  Length 9;
  Best Local Similarity   42.9%;  Pred. No. 1e+05;
  Matches    3;  Conservative    2;  Mismatches    1;  Indels    1;  Gaps    1;

Qy       5 YGLFSVKR 9
           :|:|
Db       3 NEESIGKT 7


RESULT   7
ALL7_CYDPO
ID   ALL7_CYDPO      STANDARD;      PRT;     7 AA.
AC   P82156;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   Cydiastatin 7.
OS   Cydia pomonella (Codling moth).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
OC   Tortricoidea; Tortricidae; Olethreutinae; Cydia.
OX   NCBI_TaxID=82600;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Larva;
RX   MEDLINE=98545439; PubMed=9392829;
RA   Duve H., Johnsen A.H., Maestro J.L., Scott A.G., Whitehead I.;
RA   Davey M., East P.D., Thorpe A.;
RT   "Lepidopteran peptides of the allatostatin superfamily.";
RL   Peptides 18:1301-1309(1997).
CC   -!- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.
KW   Neuropeptide; Amidation.
FT   MOD_RES       7       7       AMIDATION.
SQ   SEQUENCE   7 AA;  [illegible] MW;  [illegible];


  Query Match             28.4%;  Score 13;  DB 1;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 1e+05;
  Matches    2;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 YD 2
           ||
Db       4 YD 5


RESULT   8
ALL5_CALVO
ID   ALL5_CALVO      STANDARD;      PRT;     8 AA.
AC   P41841;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Callatostatin 5 (Met-callatostatin 1) ([Hyp3]Met callatostatin).
OS   Calliphora vomitoria (Blue bottle fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Oestroidea; Calliphoridae; Calliphora.
OX   NCBI_TaxID=7454;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=93211980; PubMed=8460157;
RA   Duve H., Johnsen A.H., Scott A.G., Yu C.G., Yazdi K., Thorpe S.S.,
RA   Thorpe A.;
RT   "Callatostatins: neuropeptides from the blowfly Calliphora vomitoria
RT   with sequence homology to cockroach allatostatins.";
RL   Proc. Natl. Acad. Sci. U.S.A. 90:2456-2460(1993).
RN   [2]
RP   CHARACTERIZATION, AND HYDROXYLATION.
RC   TISSUE=Head;
RX   MEDLINE=94342269; PubMed=8063725;
RA   Duve H., Johnsen A.H., Scott A.G., East P., Thorpe A.;
RT   "[Hyp3]Met callatostatin. Identification and biological properties of
RT   a novel neuropeptide from the blowfly Calliphora vomitoria.";
RL   J. Biol. Chem. 269:21059-21066(1994).
CC   -!- FUNCTION: MAY ACT AS A NEUROTRANSMITTER, A NEUROMODULATOR AND MAY
CC       ALSO BE INVOLVED IN THE INTEGRATION OF THE SMELL SYSTEM WITHIN THE BRAIN. MAY BE
CC       INVOLVED IN THE CONTROL OF VISCERAL MUSCLES [illegible] IT IS ABILITY TO
CC       BEHAVE AS POTENT INHIBITORS OF PERISTALTIC MOVEMENTS. MAY ALSO
CC       FULFILL A NEUROHORMONAL ROLE ON MUSCLE OF THE GUT AND HEART.
CC   -!- TISSUE SPECIFICITY: NEURONS WITHIN BRAIN AND ABDOMINAL GANGLION.
CC   -!- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.
DR   PIR; B47393; B47393.
KW   Neuropeptide; Amidation; Hydroxylation.
FT   MOD_RES       3       3       HYDROXYLATION (IN [illegible] OF THE [illegible]).
FT   MOD_RES       9       9       AMIDATION.
SQ   SEQUENCE   8 AA;  868 MW;  7D9879CAB647[illegible] 86C64;


  Query Match             28.4%;  Score 13;  DB 1;  Length 8;
  Best Local Similarity   100.0%;  Pred. No. 1e+05;
  Matches    2;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 YD 2
           ||
Db       4 YD 5


RESULT   9
BIOA_CITFR
ID   BIOA_CITFR      STANDARD;      PRT;     5 AA.
AC   P13071;
DT   01-JAN-1990 (Rel. 13, Created)
DT   01-JAN-1990 (Rel. 13, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Adenosylmethionine-8-amino-7-oxononanoate aminotransferase
DE   (EC 2.6.1.62) (7,8-diaminopelargonic acid aminotransferase) (DAPA
DE   aminotransferase) (Fragment).
GN   BIOA.
OS   Citrobacter freundii.
```

```
CC   Belongs to the class-I pyridoxal-phosphate-dependent aminotransferase family;
CC   ' ' subfamily.
CC   NOTE: LYSINE 6.
SQ
DR   SIMILARITY TO NUCLEIC ACID SEQUENCE.
DR   MEDLINE; 84161822; PubMed 6423621.
RA   Nakamura S., Takenaka O., Takahashi K.;
RT   "Fibrinopeptides A and B of baboons (Papio anubis, Papio hamadryas,
RT   and Theropithecus gelada): their amino acid sequences and
RT   evolutionary rates and a molecular phylogeny for the baboons.";
RL   J. Biochem. 94:1973-1978(1983).
CC   -!- FUNCTION: FIBRINOGEN HAS A DOUBLE FUNCTION: YIELDING MONOMERS THAT
CC       POLYMERIZE INTO FIBRIN AND ACTING AS A COFACTOR IN PLATELET
CC       AGGREGATION.
CC   -!- SUBUNIT: HEXAMER CONTAINING 2 SETS OF 3 NONIDENTICAL CHAINS
CC       (ALPHA, BETA AND GAMMA), LINKED TO EACH OTHER BY DISULFIDE BONDS.
CC   -!- MISCELLANEOUS: CONVERSION OF FIBRINOGEN TO FIBRIN IS TRIGGERED BY
CC       THROMBIN, WHICH CLEAVES FIBRINOPEPTIDES A AND B FROM ALPHA & BETA
CC       CHAINS, AND THUS EXPOSES THE N-TERMINAL POLYMERIZATION SITES
CC       RESPONSIBLE FOR THE FORMATION OF THE SOFT CLOT.
DR   PIR; D28854; D28854.
DR   InterPro; IPR002181; Fibrinogen_C.
DR   PROSITE; PS00514; FIBRIN_AG_C_DOMAIN; PARTIAL.
KW   Blood coagulation; Plasma.
FT   PEPTIDE         1       9       FIBRINOPEPTIDE B.
FT   NON_TER         9       9
SQ   SEQUENCE   9 AA;  1076 MW;  DEFE6408C7287B06 CRC64;


  Query Match             26.1%;  Score 12;  DB 1;  Length 9;
  Best Local Similarity   40.0%;  Pred. No. 1e+05;
  Matches    2;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       3 DQESV 7
           :|| :
Db       1 NQEGL 5


RESULT  11
FIBB_PAPHA
ID   FIBB_PAPHA     STANDARD;      PRT;     9 AA.
AC   P19343;
DT   01-NOV-1990 (Rel. 16, Created)
DT   01-NOV-1990 (Rel. 16, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
DE   Fibrinogen beta chain [Contains: Fibrinopeptide B] (Fragment).
GN   FGB.
OS   Papio hamadryas (Hamadryas baboon).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
OC   Cercopithecinae; Papio.
OX   NCBI_TaxID=9557;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=84161822; PubMed=6423621;
RA   Nakamura S., Takenaka O., Takahashi K.;
RT   "Fibrinopeptides A and B of baboons (Papio anubis, Papio hamadryas,
RT   and Theropithecus gelada): their amino acid sequences and
RT   evolutionary rates and a molecular phylogeny for the baboons.";
RL   J. Biochem. 94:1973-1978(1983).
CC   -!- FUNCTION: FIBRINOGEN HAS A DOUBLE FUNCTION: YIELDING MONOMERS THAT
CC       POLYMERIZE INTO FIBRIN AND ACTING AS A COFACTOR IN PLATELET
CC       AGGREGATION.
CC   -!- SUBUNIT: HEXAMER CONTAINING 2 SETS OF 3 NONIDENTICAL CHAINS
CC       (ALPHA, BETA AND GAMMA), LINKED TO EACH OTHER BY DISULFIDE BONDS.
CC   -!- MISCELLANEOUS: CONVERSION OF FIBRINOGEN TO FIBRIN IS TRIGGERED BY
CC       THROMBIN, WHICH CLEAVES FIBRINOPEPTIDES A AND B FROM ALPHA & BETA
CC       CHAINS, AND THUS EXPOSES THE N-TERMINAL POLYMERIZATION SITES
CC       RESPONSIBLE FOR THE FORMATION OF THE SOFT CLOT.
DR   PIR; E28854; E28854.
DR   InterPro; IPR002181; Fibrinogen_C.
DR   PROSITE; PS00514; FIBRIN_AG_C_DOMAIN; PARTIAL.
KW   Blood coagulation; Plasma.
FT   PEPTIDE         1       9       FIBRINOPEPTIDE B.
FT   NON_TER         9       9
SQ   SEQUENCE   9 AA;  1057 MW;  DDFE71E9C7287B06 CRC64;


  Query Match             26.1%;  Score 12;  DB 1;  Length 9;
  Best Local Similarity   40.0%;  Pred. No. 1e+05;
  Matches    2;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       3 DQESV 7
           :|| :
Db       1 NQEGL 5


RESULT  12
FIBB_THEGE
ID   FIBB_THEGE     STANDARD;      PRT;     9 AA.
```

```
AC   P19844;
DT   01-NOV 1990 (Rel. 16, Created)
DT   01-NOV 1990 (Rel. 16, Last sequence update)
DT   31-MAR 2002 (Rel. 41, Last annotation update)
DE   Fibrinogen beta chain [Contains: Fibrinopeptide B] (Fragment).
GN   FIB.
OS   Theropithecus gelada (Gelada baboon).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
OC   Cercopithecinae; Theropithecus.
OX   NCBI_TaxID=9565;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=84154822; PubMed=6423621;
RA   Nakamura S., Takeuchi Y., Takagi T.;
RT   "Fibrinopeptides A and B of baboons (Papio anubis, Papio hamadryas,
RT   and Theropithecus gelada): their amino acid sequences and
RT   evolutionary rates and a molecular phylogeny for the baboons.";
RL   J. Biochem. 94:1973-1978(1983).
CC   -!- FUNCTION: FIBRINOGEN HAS A DOUBLE FUNCTION. YIELDING MONOMERS THAT
CC       POLYMERIZE INTO FIBRIN AND ACTING AS A COFACTOR IN PLATELET
CC       AGGREGATION.
CC   -!- SUBUNIT: HEXAMER CONTAINING 2 SETS OF 3 NONIDENTICAL CHAINS
CC       (ALPHA, BETA AND GAMMA), LINKED TO EACH OTHER BY DISULFIDE BONDS.
CC   -!- MISCELLANEOUS: CONVERSION OF FIBRIN TO FIBRIN POLYMER IS
CC       THROMBIN, WHICH CLEAVES FIBRINOPEPTIDES A AND B FROM ALPHA & BETA
CC       CHAINS, AND THUS EXPOSES THE N TERMINAL POLYMERIZATION SITES
CC       RESPONSIBLE FOR THE FORMATION OF THE SOFT CLOT.
DR   PIR; A28854; A28854.
DR   InterPro; IPR002181; Fibrinogen_C.
DR   PROSITE; PS00514; FIBRINOGEN_C_DOMAIN; PARTIAL.
KW   Blood coagulation; Plasma.
FT   PEPTIDE        1      9       FIBRINOPEPTIDE B
FT   NON_TER        9      9
SQ   SEQUENCE   9 AA;  977 MW;  DDFE7879C7287B05 CRC64;


  Query Match          26.1%;  Score 12;  DB 1;  Length 9;
  Best Local Similarity   40.0%;  Pred. No. 1e+05;
  Matches    2;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       3 DQESV 7
           : | :
Db       1 NQEGL 5


RESULT  13
MOSE_CLYJA
ID   MOSE_CLYJA     STANDARD;      PRT:     9 AA.
AC   P19845;
DT   01-FEB 1991 (Rel. 17, Created)
DT   01-FEB 1991 (Rel. 17, Last sequence update)
DT   01-FEB 1991 (Rel. 17, Last annotation update)
DE   [Glu-6] mosaet.
OS   Clypeaster japonicus (Sand dollar).
OC   Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
OC   Echinoidea; Euechinoidea; Gnathostomata; Clypeasteroida;
OC   Clypeasteridae; Clypeaster.
OX   NCBI_TaxID=7644;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Egg jelly;
RA   Suzuki N., Fujita M., Yoshino K.I., Kajiura H., Nomura K.;
RA   Yamaguchi M.;
RT   "Purification and structure of mosaet and its derivatives from the
RT   egg jelly of the sea urchin Clypeaster japonicus.";
RL   Zool. Sci. 4:649-656(1987).
CC   -!- FUNCTION: STIMULATES SPERM RESPIRATION AND MOTILITY.
DR   PIR; JN0027; JN0027.
SQ   SEQUENCE   9 AA;  924 MW;  93245729DC5BAB5 CRC64;


  Query Match          26.1%;  Score 12;  DB 1;  Length 9;
  Best Local Similarity   50.0%;  Pred. No. 1e+05;
  Matches    2;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       3 DQES 6
           | :|
Db       1 DSDS 4


RESULT  14
MOSB_CLYJA
ID   MOSB_CLYJA     STANDARD;      PRT:     9 AA.
AC   P19852;
DT   01-FEB 1991 (Rel. 17, Created)
DT   01-FEB 1991 (Rel. 17, Last sequence update)
DT   01-FEB 1991 (Rel. 17, Last annotation update)
DE   [His-6] mosaet.
OS   Clypeaster japonicus (Sand dollar).
OC   Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
OC   Echinoidea; Euechinoidea; Gnathostomata; Clypeasteroida;
OC   Clypeasteridae; Clypeaster.
OX   NCBI_TaxID=7644;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Egg jelly;
RA   Suzuki N., Fujita M., Yoshino K.I., Kajiura H., Nomura K.;
RA   Yamaguchi M.;
RT   "Purification and structure of mosaet and its derivatives from the
RT   egg jelly of the sea urchin Clypeaster japonicus.";
RL   Zool. Sci. 4:649-656(1987).
CC   -!- FUNCTION: STIMULATES SPERM RESPIRATION AND MOTILITY.
DR   PIR; JN0026; JN0026.
SQ   SEQUENCE   9 AA;  914 MW;  93245729EEAC5BAB5 CRC64;


  Query Match          26.1%;  Score 12;  DB 1;  Length 9;
  Best Local Similarity   50.0%;  Pred. No. 1e+05;
  Matches    2;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       3 DQES 6
           | :|
Db       1 DSDS 4


RESULT  15
UHAJ_HUMAN
ID   UHAJ_HUMAN     STANDARD;      PRT:     9 AA.
AC   P40929;
DT   01-FEB 1995 (Rel. 31, Created)
DT   01-FEB-1995 (Rel. 31, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Unknown protein from 2D-page of heart (Spot 5613)  Fragment).
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Heart;
RX   MEDLINE=95204287; PubMed=7895732;
RA   Corbett J.M., Wheeler C.H., Baker C.S., Yacoub M.H., Dunn M.J.;
RT   "The human myocardial two-dimensional gel protein database: update
RT   1994.";
RL   Electrophoresis 15:1459-1465(1994).
CC   -!- MISCELLANEOUS: ON THE 2D GEL THE DETERMINED PI OF THIS UNKNOWN
CC       PROTEIN IS: 6.0, ITS MW IS: 55.3 KDa.
FT   NON_TER        9      9
SQ   SEQUENCE   9 AA;  1104 MW;  8874B1BB7B0C8A7A CRC64;


  Query Match          26.1%;  Score 12;  DB 1;  Length 9;
  Best Local Similarity   66.7%;  Pred. No. 1e+05;
  Matches    2;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;
```

GeneCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 26, 2002, 06:39:45 ; Search time 24.24 Seconds
(without alignments)
11.455 Million cell updates/sec

Title: US-09-744-282-3
Perfect score: 49
Sequence: 1 YDDQE 5

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 556222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 556222

Minimum DB seq length: 0
Maximum DB seq length: 5

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_19:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result in this printout, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	49	84.8	8	4	Q9UJ50	Q9uj50 homo sapien
2	36	84.8	8	11	Q9QVD3	Q9qvd3 rattus sp.
3	35	82.6	9	8	Q9I2K9	Q9i2k9 spinacia ol
4	25	52.6	9	10	P82440	P82440 triticum a
5	24	50.4	7	10	O49223	O49223 glycine max
6	24	50.4	8	6	Q9BEA7	Q9bea7 macroscelid
7	24	50.4	9	2	P70740	P70740 alcaligenes
8	24	50.4	9	12	Q82622	Q82622 urban intes
9	24	28.3	7	6	Q28742	Q28742 oryctolagus
10	24	28.3	8	2	P83158	P83158 arabaena sp
11	24	28.3	8	4	Q96Q65	Q96q65 homo sapien
12	24	28.3	9	4	Q9UCQ9	Q9ucq9 homo sapien
13	24	28.3	9	4	Q9BQT4	Q9bqt4 homo sapien
14	24	28.3	9	5	P92292	P92292 t.thyx mrii
15	24	26.1	8	2	Q9R8X0	Q9r8x0 planktothri
16	24	26.1	8	4	Q9H4I6	Q9h4i6 homo sapien
17	24	26.1	8	5	Q9N0M5	Q9n0m5 toxoplasma
18	24	26.1	8	13	Q9PS69	Q9ps69 gallus gall
19	24	26.1	9	4	Q9UK44	Q9uk44 homo sapien
20	24	26.1	9	4	Q96KF6	Q96kf6 homo sapien
21	24	26.1	9	6	Q28095	Q28095 bos taurus
22	24	26.1	9	6	P82926	P82926 bos taurus
23	24	26.1	9	12	Q91IX6	Q91ix6 marinera hemo
24	24	23.9	7	2	Q47477	Q47477 ostertagia
25	24	23.9	8	2	P72279	P72279 rhodococcus
26	24	23.9	8	5	Q94695	Q94695 physarum po
27	24	23.9	8	12	Q9WI85	Q9wi85 pseudorabie
28	24	23.9	9	2	Q47556	Q47556 escherichia
29	24	23.9	9	2	Q943E0	Q943e0 streptococc
30	24	23.9	9	4	Q9BYE9	Q9bye9 homo sapien
31	24	23.9	9	4	Q14715	Q14715 homo sapien
32	24	23.9	9	4	Q96178	Q96178 homo sapien
33	24	23.9	9	6	Q9XS10	Q9xs10 capra hircu
34	10	21.7	5	2	P83073	P83073 bacillus ce
35	10	21.7	6	10	P82541	P82541 spisula so
36	10	21.7	7	14	P82161	P82161 silurus cl
37	10	21.7	8	2	Q47275	Q47275 escherichia
38	10	21.7	8	4	Q9P285	Q9p285 homo sapien
39	10	21.7	8	4	Q15895	Q15895 homo sapien
40	10	21.7	8	5	Q94623	Q94623 manduca sex
41	10	21.7	8	8	Q36898	Q36898 trichosurus
42	10	21.7	8	13	P83408	P83408 canis rhyncha
43	10	21.7	9	2	Q46179	Q46179 clostridium
44	10	21.7	9	2	P72149	P72149 pseudomonas
45	10	21.7	9	2	P83167	P83167 anabaena sp

ALIGNMENTS

RESULT 1
Q9UJ50
ID Q9UJ50 PRELIMINARY; PRT; 8 AA.
AC Q9UJ50;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE LAIROPHILIN 2 (FRAGMENT).
GN LPHH1.
OS Homo Sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99154747; PubMed=10030676;
RA White G.R.M., Varley J.M., Heighway J.;
RT "Isolation and characterisation of a 350kb human BAC-YAC tile
RT lairophilin gene from a region of 1p31.2 implicated in breast
RT cancer.";
RL Oncogene 17:4513-4519(1998).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=20225451; PubMed=11786532;
RA White G.R.M., Varley J.M., Heighway J.;
RT "Genomic structure and expression profile of LPHH1, a 7TM gene
RT variably expressed in breast cancer cell lines.";
RL Biochim. Biophys. Acta 1491:75-92(2001).
DR EMBL; AJ244509; CAB60204.1; -.
FT NON_TER 1 1
FT NON_TER 8 8
SQ SEQUENCE 8 AA; 102% MW; 4E5C6C6C 44AA4D8C CRC64;

Query Match 84.8%; Score 49; DB 4; Length 8;
Best Local Similarity 40.0%; Pred. No. 1.6e-05;
Matches 2; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 YDDQE 5

```
                                                                                
                                                                                
RESULT  2
Q9LKG9
ID   Q9LKG9      PRELIMINARY;      PRT:     9 AA.
AC   Q9LKG9;
DT   01-MAY-2000 (TrEMBLrel. 14, Created)
DT   01-MAY-2000 (TrEMBLrel. 14, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 14, Last annotation update)
DE   PSII 11 KDA LIGHT HARVESTING CHLOROPHYLL PROTEIN CP.
OS   Spinacia oleracea (Spinach).
OC   Chloroplast.
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC   Caryophyllales; Chenopodiaceae; Spinacia.
OX   NCBI_TaxID=3562;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=99367679; PubMed=10394641;
RA   Ma Y., Hall H., Brittin K.R., Shabanowitz J., Hunt D.F., Bassett C.;
RT   "Tandem mass spectrometry identifies sites of three post-translational
RT   modifications of spinach light harvesting chlorophyll protein CP.
RT   29: cleavage, acetylation, and phosphorylation.";
RL   J. Prot. Chem. 26:175-184 1791(1999).
SQ   SEQUENCE   9 AA;  1042 MW;  6B5G6D8C5B322C1B4 CRC64;


  Query Match             32.6%;  Score 15;  DB 2;  Length 9;
  Best Local Similarity   50.0%;  Pred. No. 5.6e+05;
  Matches    3;  Conservative    2;  Mismatches    1;  Indels    1;  Gaps    0;

Qy       1 YEVK 4
            |:|:
Db       8 WEVK 5


RESULT  3
Q9LKG8
ID   Q9LKG8      PRELIMINARY;      PRT:     9 AA.
AC   Q9LKG8;
DT   01-MAY-2000 (TrEMBLrel. 14, Created)
DT   01-MAY-2000 (TrEMBLrel. 14, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 14, Last annotation update)
DE   PSII 11 KDA LIGHT HARVESTING CHLOROPHYLL PROTEIN CP.
OS   Spinacia oleracea (Spinach).
OC   Chloroplast.
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC   Caryophyllales; Chenopodiaceae; Spinacia.
OX   NCBI_TaxID=3562;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=99367679; PubMed=10394641;
RA   Ma Y., Hall H., Brittin K.R., Shabanowitz J., Hunt D.F., Bassett C.;
RT   "Tandem mass spectrometry identifies sites of three post-translational
RT   modifications of spinach light harvesting chlorophyll protein CP.
RT   29: cleavage, acetylation, and phosphorylation.";
RL   J. Prot. Chem. 26:175-184 1791(1999).
SQ   SEQUENCE   9 AA;  1042 MW;  6B5G6D8C5B322C1B4 CRC64;


  Query Match             32.6%;  Score 15;  DB 2;  Length 9;
  Best Local Similarity   50.0%;  Pred. No. 5.6e+05;
  Matches    3;  Conservative    2;  Mismatches    1;  Indels    1;  Gaps    0;

Qy       1 YEVK 4
            |:|:
Db       8 WEVK 5


RESULT  4
Q9LKG7
ID   Q9LKG7      PRELIMINARY;      PRT:     9 AA.
AC   Q9LKG7;
DT   01-MAY-2000 (TrEMBLrel. 14, Created)
DT   01-MAY-2000 (TrEMBLrel. 14, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 14, Last annotation update)
DE   PSII 11 KDA LIGHT HARVESTING CHLOROPHYLL PROTEIN CP.
OS   Spinacia oleracea (Spinach).
OC   Chloroplast.
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC   Caryophyllales; Chenopodiaceae; Spinacia.
OX   NCBI_TaxID=3562;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=99367679; PubMed=10394641;
RA   Ma Y., Hall H., Brittin K.R., Shabanowitz J., Hunt D.F., Bassett C.;
RT   "Tandem mass spectrometry identifies sites of three post-translational
RT   modifications of spinach light harvesting chlorophyll protein CP.
RT   29: cleavage, acetylation, and phosphorylation.";
RL   J. Prot. Chem. 26:175-184 1791(1999).
SQ   SEQUENCE   9 AA;  1042 MW;  6B5G6D8C5B322C1B4 CRC64;


  Query Match             32.6%;  Score 15;  DB 2;  Length 9;
  Best Local Similarity   50.0%;  Pred. No. 5.6e+05;
  Matches    3;  Conservative    2;  Mismatches    1;  Indels    1;  Gaps    0;

Qy       1 YEVK 4
            |:|:
Db       8 WEVK 5


RESULT  5
Q9LKG6
ID   Q9LKG6      PRELIMINARY;      PRT:     9 AA.
AC   Q9LKG6;
DT   01-MAY-2000 (TrEMBLrel. 14, Created)
DT   01-MAY-2000 (TrEMBLrel. 14, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 14, Last annotation update)
DE   PSII 11 KDA LIGHT HARVESTING CHLOROPHYLL PROTEIN CP.
OS   Spinacia oleracea (Spinach).
OC   Chloroplast.
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC   Caryophyllales; Chenopodiaceae; Spinacia.
OX   NCBI_TaxID=3562;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=99367679; PubMed=10394641;
RA   Ma Y., Hall H., Brittin K.R., Shabanowitz J., Hunt D.F., Bassett C.;
RT   "Tandem mass spectrometry identifies sites of three post-translational
RT   modifications of spinach light harvesting chlorophyll protein CP.
RT   29: cleavage, acetylation, and phosphorylation.";
RL   J. Prot. Chem. 26:175-184 1791(1999).
SQ   SEQUENCE   9 AA;  1042 MW;  6B5G6D8C5B322C1B4 CRC64;


  Query Match             32.6%;  Score 15;  DB 2;  Length 9;
  Best Local Similarity   50.0%;  Pred. No. 5.6e+05;
  Matches    3;  Conservative    2;  Mismatches    1;  Indels    1;  Gaps    0;

Qy       1 YEVK 4
            |:|:
Db       8 WEVK 5


RESULT  6
P82440
ID   P82440      PRELIMINARY;      PRT:     9 AA.
AC   P82440;
DT   01-JUN-2000 (TrEMBLrel. 14, Created)
DT   01-JUN-2000 (TrEMBLrel. 14, Last sequence update)
DT   01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE   42 KDA CELL WALL PROTEIN (FRAGMENT).
OS   Nicotiana tabacum (Common tobacco).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC   Asteridae; euasterids 1; Solanales; Solanaceae; Nicotiana.
OX   NCBI_TaxID=4097;
RN   [1]
RP   SEQUENCE.
RC   STRAIN=CV. PETIT HAVANA;
RA   Blee K.A., Bonham V.A., Mitchell G.P., Robertson D., Stabis A.R.,
RA   Wojtaszek P., Bolwell G.P.;
RT   "Proteomic study of secondary cell wall proteins from transformed
RT   tobacco culture.";
RL   Planta 0:0-0(2000).
CC   -!- SUBCELLULAR LOCATION: CELL WALL.
CC   -!- TISSUE SPECIFICITY: XYLEM.
KW   Cell wall.
FT   NON_TER       9       9
SQ   SEQUENCE   9 AA;  1053 MW;  29BCC9D2D5B81B07 CRC64;


  Query Match             32.6%;  Score 15;  DB 1;  Length 9;
  Best Local Similarity   75.0%;  Pred. No. 5.6e+05;
  Matches    3;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       4 QESV 7
            :|||
Db       4 EESV 6


RESULT  7
O49223
ID   O49223      PRELIMINARY;      PRT:     7 AA.
AC   O49223;
DT   01-JUN-1998 (TrEMBLrel. 06, Created)
DT   01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   HMG-I-LIKE PROTEIN (FRAGMENT).
OS   Glycine max (Soybean).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
OX   NCBI_TaxID=3847;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CV. ESSEX; TISSUE=ROOT;
RX   MEDLINE=91367679; PubMed=1891369;
RA   Laux L., Goldberg R.B.;
RT   "A plant DNA binding protein shares highly conserved sequence motifs
RT   with HMG-box proteins.";
RL   Nucleic Acids Res. 19:4769-4769(1991).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CV. ESSEX; TISSUE=ROOT;
RA   Mahalingam R., Knap H.T.;
RL   Submitted (FEB-1998) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF047050; AAC04556.1; -.
FT   NON_TER       1       1
SQ   SEQUENCE   7 AA;  850 MW;  6AAAAAB47863781D CRC64;


  Query Match             30.4%;  Score 14;  DB 1;  Length 7;
  Best Local Similarity   66.7%;  Pred. No. 5.3e+05;
  Matches    2;  Conservative    1;  Mismatches    0;  Indels    1;  Gaps    0;

Qy       3 YEVK 6
            :|||
Db       4 WEVK 7
```

RESULT 6
A_HBFA'
ID A_HBFA' PRELIMINARY; PRT; 8 AA.
AC A_HBFA';
DT [illegible] (1:EMBL:rel. 17, Created)
DT [illegible] (1:EMBL:rel. 17, last sequence update)
DT [illegible] (1:EMBL:rel. 17, last annotation update)
DE 'AME BIFF NCIVE ELEMENT MESENAT E (FRAGMENT).
GN 'FEM.
OS Macroscelides proboscideus (short eared elephant shrew).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Macroscelidea; Macroscelididae; Macroscelides.
OX NCBI_TaxID=27662;
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE [illegible]; PubMed 11214319;
RA Murray W.J., Fichtel E., Johnson W.E., Chan Y.F., [illegible] S.J.;
RA Bell S.J.;
RT "Molecular phylogenetics and the origins of placental mammals.";
RL Nature 409:614-618(2001).
DR EMBL; AY015695; AAM47551.1; -.
FT NON_TER 1 1
SQ SEQUENCE 9 AA; 1113 MW; [illegible] CRC64;

Query Match 50.4%; Score 14; DB 2; Length 9;
Best Local Similarity 50.0%; Pred. No. 5.6e+05;
Matches 2; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 YDQ 4
| :|
Db 8 YSDE 5

RESULT 7
P70340
ID P70340 PRELIMINARY; PRT; 9 AA.
AC P70340;
DT 01-FEB-1997 (TrEMBLrel. 02, Created)
DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE HOXA PROTEIN (FRAGMENT).
GN HOXA.
OS Alcaligenes eutrophus (Ralstonia eutropha).
OC Bacteria; Proteobacteria; beta subdivision; Ralstonia group;
OC Ralstonia.
OX NCBI_TaxID=519;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN H16.
RX MEDLINE 95124179; PubMed 8021224;
RA Lenz O., Schwartz E., Dernedde J., Eitinger M., Friedrich B.;
RT "The Alcaligenes eutrophus H16 hoxX gene participates in hydrogenase
RT regulation.";
RL J. Bacteriol. 176:4385-4393(1994).
DR EMBL; X74067; CAA52746.1; -.
FT NON_TER 9 9
SQ SEQUENCE 9 AA; 994 MW; 970B01AE6D33AB5 CRC64;

Query Match 50.4%; Score 14; DB 2; Length 9;
Best Local Similarity 50.0%; Pred. No. 5.6e+05;
Matches 2; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 2 PDQESV 7
| :|
Db 4 PEKAIV 9

RESULT 8
Q82622
ID [illegible] PRELIMINARY; PRT; 9 AA.
AC Q82622;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE ORF A (FRAGMENT).
OS Avian infectious bronchitis virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
OC Coronaviridae; Coronavirus.
OX NCBI_TaxID=11120;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE 86010264; PubMed 2995560;
RA Boursnell M.E., Binns M.M., Brown T.D.S.;
RT "Sequencing of coronavirus IBV genomic RNA: three open reading frames
RT in the 5' 'unique' region of mRNA 1.";
RL J. Gen. Virol. 66:2253-2258(1985).
DR EMBL; M25338; AAA46234.1; -.
FT NON_TER 1 1
SQ SEQUENCE 9 AA; 1134 MW; [illegible] CRC64;

Query Match 50.4%; Score 14; DB 12; Length 9;
Best Local Similarity 42.9%; Pred. No. 5.6e+05;
Matches 3; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 YDQESV 7
| : ::|
Db 3 YRPKKSV 9

RESULT 9
Q28742
ID Q28742 PRELIMINARY; PRT; 7 AA.
AC Q28742;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE ALPHA MYOSIN HEAVY CHAIN (FRAGMENT).
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Laurasiatheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE 84221901; PubMed 6428491;
RA Friedman D.J., Umeda P.K., Sinha A.M., Bou H. J., Kovacs S.,
RA Rabinowitz M.;
RT "Characterization of genomic clones specify. a rabbit alpha- and beta
RT ventricular myosin heavy chains.";
RL Proc. Natl. Acad. Sci. U.S.A. 81:3944-3948(1984).
DR EMBL; K01698; AAA31415.1; -.
FT NON_TER 1 1
SQ SEQUENCE 7 AA; 916 MW; [illegible] CRC64;

Query Match 28.4%; Score 8; DB 2; Length 7;
Best Local Similarity 66.7%; Pred. No. 5.6e+05;
Matches 2; Conservative 1; Mismatches 0; Indels 1; Gaps 0;

Qy 3 DQE 5
| :|
Db 5 DEE 7

RESULT 10
P83158
ID P83158 PRELIMINARY; PRT; 9 AA
AC P83158;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE PHOTOSYSTEM I IRON-SULFUR CENTER (PHOTOSYSTEM I SUBUNIT VII) (P-KDA
DE POLYPEPTIDE) (PSI C) (FRAGMENT).

OS Anabaena sp. (strain L31).
OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Anabaena.
OX NCBI_TaxID=264;
RN [1]
RP SEQUENCE.
RA Apte S.K., Straus D.B., Schmid R., Altendorf K.;
RL Submitted (OCT-2001) to the SWISS-PROT data bank.
CC -!- FUNCTION: AP-PROTEIN FOR THE IRON SULFUR CENTER, FA AND FB OF THE
CC PHOTOSYSTEM I COMPLEX.
CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC -!- SIMILARITY: THE IRON SULFUR CENTERS ARE SIMILAR TO THE SE-F
CC BACTERIAL TYPE 4FE-4S FERREDOXINS.
DR EMBL; AB056340; BAB54936.1; -.
DR PIR; T10747; T10747.
KW Iron-sulfur protein; Photosystem I; Iron-sulfur; 4Fe-4S; Metal-binding.
FT NON_TER 9 9
SQ SEQUENCE 9 AA; 962 MW; C5BB50532DDA1F5 CRC64;

Query Match 28.3%; Score 13; DB 2; Length 9;
Best Local Similarity 100.0%; Pred. No. 5.6e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YD 2
||
Db 1 YD 2

RESULT 13
Q9BQI4
ID Q9BQI4 PRELIMINARY; PRT; 9 AA.
AC Q9BQI4;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE CLIP-ASSOCIATING PROTEIN CLASP2 (FRAGMENT).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BRAIN;
RA Galjart N.;
RL Submitted (MAR-2001) to the EMBL/GenBank/DDBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=BRAIN;
RX MEDLINE=21165938; PubMed=11290329;
RA Akhmanova A., Hoogenraad C.C., Drabek K., Stepanova T., Dortland B.,
RA Verkerk T., Vermeulen W., Burgering B.M., de Zeeuw C.I., Grosveld F.,
RA Galjart N.;
RT "CLASPs are CLIP-115 and -170 associating proteins involved in the
RT regional regulation of microtubule dynamics in motile fibroblasts.";
RL Cell 104:923-935(2001).
DR EMBL; AJ288058; CAC35157.1; -.
FT NON_TER 9 9
SQ SEQUENCE 9 AA; 1001 MW; C5E8E63AAAA87C9D CRC64;

Query Match 28.3%; Score 13; DB 4; Length 9;
Best Local Similarity 66.7%; Pred. No. 5.6e+05;
Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 HDQ 4
|::
Db 5 HDK 7

RESULT 14
P82208
ID P82208 PRELIMINARY; PRT; 9 AA.
AC P82208;
DT 01-OCT-2001 (TrEMBLrel. 18, Created)
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT 01-OCT-2001 (TrEMBLrel. 18, Last annotation update)
DE UNKNOWN PROTEIN FROM 2D PAGE (FRAGMENT).

OS Bombyx mori (Silk moth).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
OC Bombycoidea; Bombycidae; Bombyx.
OX NCBI_TaxID=7091;
RN [1]
RP SEQUENCE.
RC STRAIN=XINGANG X KEMENG, FIFTH BODY WALL AND FAT BODY;
RX MEDLINE=21274481; PubMed=11280994;
RA Zhu J.X.;
RT "Protein database for several tissues derived from the silkworm.";
RL J Chuan Hsueh Pao 28:217-224(2001).
FT NON_TER 1 1
SQ SEQUENCE 1 AA; 1009 MW; 877C3ABB45B042D7 CRC64;

Query Match 28.4%; Score 13; DB 5; Length 1;
Best Local Similarity 100.0%; Prod. No. 5.6e-05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YI 2
:: ::
Db 1 YI 2

RESULT 2
Q9BSX0
ID Q9BSX0 PRELIMINARY; PRT; 8 AA.
AC Q9BSX0;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE RIBOSOMAL SPECIFIC ENDONUCLEASE/EXORIBONUCLEASE (ARCH. STRAIN).
DE (FRAGMENT).
GN RRNL.
OS Planktothrix rubescens.
OC Bacteria; Cyanobacteria; Oscillatoriales; Planktothrix.
OX NCBI_TaxID=67122;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CCP PLA 9516, AND CCP-PLA 9304;
RX MEDLINE=20558580; PubMed=10537192;
RA Beard S.J., Handley P.A., Hayes P.K., Walsby A.E.;
RT "The diversity of gas vesicle genes in Planktothrix rubescens in a
RT lake ecosystem.";
RL Microbiology 145:2757-2768(1999).
DR EMBL; AF154240; AAD50537.1; -.
DR EMBL; AF154248; AAD50534.1; -.
FT NON_TER 1 1
SQ SEQUENCE 8 AA; 957 MW; 43D1AAA685B81C0B CRC64;

Query Match 28.4%; Score 13; DB 2; Length 8;
Best Local Similarity 14.3%; Prod. No. 5.6e-05;
Matches 2; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

Qy 1 YIGKESV 7
:: ::
Db 2 FESMCIV 8

Search completed: August 26, 2002, 06:39:47
Job time: 9.8 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 26, 2002, 06:27:08 ; Search time 126.49 Seconds
(without alignments)
7.900 Million cell updates/sec

Title: US-09-744-282-3
Perfect score: 46
Sequence: 1 YIPQPESVKS 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 347574 seqs, 111077396 residues

Total number of hits satisfying chosen parameters: [illegible]

Minimum DB seq length: 0
Maximum DB seq length: 9

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_032802:*
1: [illegible]*
2: [illegible]*
3: [illegible]*
4: [illegible]*
5: [illegible]*
6: [illegible]*
7: [illegible]*
8: [illegible]*
9: [illegible]*
10: [illegible]*
11: [illegible]*
12: [illegible]*
13: [illegible]*
14: [illegible]*
15: [illegible]*
16: [illegible]*
17: [illegible]*
18: [illegible]*
19: [illegible]*
20: [illegible]*
21: [illegible]*
22: [illegible]*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	46	100.0	9	17	AAR95073	Human cartilage gl
2	46	100.0	9	21	AAY69228	Peptide fragment o
3	25	54.3	7	22	AAR82911	Mycobacter VIBE BF
4	24	52.2	8	10	AAP92282	Linkage peptide to
5	24	52.2	8	17	AAR89929	Linker 3 peptide 1b
6	24	52.2	8	18	AAW82245	Linker peptide 4 o
7	24	52.2	8	18	AAW81561	Linker sequence 1
8	23	50.0	6	20	AAY43571	Scrambled peptide
9	23	50.0	6	20	AAY23962	Scrambled filament
10	23	50.0	6	21	AAY67529	Scrambled FHA pept
11	23	50.0	9	22	AAU04448	Neisseria meningit
12	22	47.8	9	15	AAR69861	Peptide (26) indi
13	22	47.8	9	20	AAY43632	Immunogenic peptid
14	21	45.7	6	9	AAP82908	Peptide conju 248
15	21	45.7	7	15	AAR43052	HLA restricted T h
16	21	45.7	7	16	AAR80465	Japanese cystic fi
17	21	45.7	7	22	AAR48174	HLA epitope used
18	21	45.7	8	17	AAR93035	Linear chain acid
19	21	45.7	8	21	AAY77718	Xpress peptide epi
20	21	45.7	8	22	AAR97365	Xpress epitope seq
21	21	45.7	8	22	AAR98869	Xpress peptide epi
22	21	45.7	9	15	AAR57724	Peptide (224) indi
23	21	45.7	9	20	AAY41264	TOR peptide derivv
24	20	43.5	7	18	AAW29967	Cysteine protease
25	20	43.5	8	10	AAP92283	Linkage peptide 2c
26	20	43.5	8	11	AAR36434	Peptide bridge Y1
27	20	43.5	8	17	AAR88039	Linker sequence 1a
28	20	43.5	8	17	AAR94035	Linear chain acet
29	20	43.5	8	18	AAW82242	Linker peptide 1 c
30	20	43.5	8	21	AAB13095	Combinatorial chel
31	20	43.5	9	20	AAY56204	Neutrophil activat
32	20	43.5	9	20	AAW97981	Human synaptotagmal
33	20	43.5	9	20	AAW97943	Human synaptotagmi
34	20	43.5	9	21	AAY76679	SCP-1 HLA binding
35	20	43.5	9	21	AAY77719	SCP-1 HLA binding
36	20	43.5	9	22	AAM09484	Vaccine to latent MB
37	20	43.5	9	22	AAU06547	Human MB7 melanoma
38	19	41.3	4	18	AAW08886	4 Amidobenzoyl hyd
39	19	41.3	4	21	AAR28367	Trypsin modulating
40	19	41.3	4	21	AAR96158	Trypsin modulating
41	19	41.3	4	21	AAR37919	Trypsin modulating
42	19	41.3	4	21	AAR70649	Trypsin modulating
43	19	41.3	7	15	AAR42116	Hydrophilic acetyl
44	19	41.3	7	20	AAR82147	Peptide #9 used in
45	19	41.3	8	18	AAR2938	Endothelin antagon

ALIGNMENTS

RESULT 1
AAR95073
ID AAR95073 standard; peptide; 9 AA.
XX
AC AAR95073;
XX
DT 17-JAN-1996 (first entry)
XX
DE Human cartilage glycoprotein-39 citrullination derived peptide (CIT).
XX
KW Human; cartilage; glycoprotein; gp-39; citrullination;
KW rheumatic; T-cell tolerance; modified; restrict; HLA;
KW rheumatoid arthritis; disease model; diagnosis.
XX
OS Homo sapiens.
XX
PN WO9613517-A1.
XX
PD 09-MAY-1996.
XX
PF 25-OCT-1995; 95WO-EP04201.
XX
PR 07-APR-1995; 95EP-0200886.
PR 27-OCT-1994; 94EP-0203128.
XX
PA (ASKO) AKE N.V.
XX
PI Boots AMBL, Verlinden CPM.
XX
DR WPI; 1996-239447/24.
XX
PT Peptide(s) derived from human cartilage gp-39 citrullination used
PT to induce specific T cell tolerance, esp. for treatment of
PT rheumatoid arthritis, also useful in producing arthritic animal

ID [illegible]
XX
[illegible] standard; peptide; 9 AA.
XX
[illegible] 22-MAY-2001 (first entry)
XX
DE [illegible] peptide, derived from the human cartilage gp-39 glycoprotein
CC [illegible] (HC gp-39) autoantigen, can be used to induce specific T cell
CC [illegible] tolerance to HC gp-39 in a patient with T cell mediated cartilage
CC [illegible] destruction, esp. rheumatoid arthritis (RA). It may also be used to
CC [illegible] induce RA in mice, etc. (human disease models), after diagnose the
CC [illegible] presence of activated T cells autoreactive with HC gp-39, i.e. to
CC [illegible] identify patients suitable for tolerising treatment.
CC [illegible] Preferably the peptide is given in a dose of ... [illegible] microg/kg
CC [illegible] by injection, or 10-50 microg/kg to produce a disease model. The
CC peptide has a specific tolerising effect on autoreactive T cells,
CC [illegible] leaves other components of the immune system intact, and
CC therefore should have few side effects. In an assay to determine
CC the peptide's binding to HLA-DR molecules, specifically DR4Dw4 (RA-ass
CC [illegible] to be associated with an increased risk of RA development), the
CC peptide gave a dot blot D(50) (micromolar) value of ...23 compared
CC [illegible] for the control peptide TBA(307-319)F.
XX
SQ Sequence [illegible] AA;

Query Match [illegible] 100.0%; Score 46; DB 17; Length 9;
Best Local Similarity [illegible] 100.0%; Pred. No. [illegible];
Matches [illegible] 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy [illegible]
Db [illegible]

RESULT [illegible]
AAY[illegible]
ID AAY[illegible] standard; peptide; 9 AA.
XX
AC AAY[illegible];
XX
DT [illegible] MAY 2001 (first entry)
XX
DE Peptide fragment of human cartilage gp-39 (HC gp-39) protein.
XX
KW Human cartilage gp-39 protein; HC gp-39 protein; T-cell activator;
KW [illegible]matory diseases; lymphocyte; antisen; bystander suppression;
KW antigen specific T cell tolerance; autoimmune disorder
KW [illegible]matory disease; rheumatoid arthritis;
KW [illegible]mmune inflammatory disease; autoantigen.
XX
OS Homo sapiens.
XX
PN WO[illegible]A.
XX
PD [illegible].
XX
PF [illegible];
XX
PR [illegible];
XX
PA (ALFA) AKZO NOBEL NV.
XX
PI Mikkelsen AMM, Boots AMH;
XX
DR WPI; [illegible].
XX
PT [illegible] cartilage protein or peptides for treating inflammatory diseases
PT especially rheumatoid arthritis.
XX
PS [illegible]; Page 17; 29pp; English.
XX
CC AAY[illegible] represent fragments of the human cartilage gp-39 (HC gp-39)
CC [illegible]. The peptides are T cell activators. The peptides are used for
CC [illegible] inflammatory diseases, by modulating lymphocytes reactive
CC to antigens other than HC gp-39, present in the same tissue of
CC HC gp-39. By the induction of antigen specific T cell tolerance,
CC autoimmune disorders can be treated by bystander suppression. The
CC HC gp-39 peptides are useful for preventing inflammatory diseases
CC especially rheumatoid arthritis. The fragments may also be used for
CC treating autoimmune inflammatory diseases. The HC gp-39 peptides are
CC especially beneficial in autoimmune conditions in which the
CC autoantigens associated with the condition are unknown.
XX
SQ Sequence 9 AA;

Query Match 100.0%; Score 46; DB 21; Length 9;
Best Local Similarity 100.0%; Pred. No. 6.4e-05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YDAQESVKS 9
|||||||||
Db 1 ydkdqesvks 9

RESULT 4
AAB81944
ID AAB81944 standard; Protein; 7 AA.
XX
AC AAB81944;
XX
DT 25-JUN-2001 (first entry)
XX
DE Marmoset VDRE-BP #1 derived tryptic peptide #5.
XX
KW Marmoset; vitamin D response element binding protein; VDRE-BP; lymphoma;
KW vitamin D resistance; osteoporosis; hypercalcaemia; vitamin D toxicity;
KW glucocorticoid mediated disorder; granuloma forming disease;
KW vitamin D intoxication; steroid hormone hypersecretion; gene therapy.
XX
OS Saguinus oedipus.
XX
PN WO200121649-A2.
XX
PD 29-MAR-2001.
XX
PF 20-SEP-2000; 2000WO-US25844.
XX
PR 22-SEP-1999; 99DE-0400967.
XX
PA (CEDA) CEDARS SINAI MEDICAL CENT.
XX
PI Adams JS;
XX
DR WPI; 2001-308082/32.
XX
PT New vitamin D response element binding protein (VDRE-BP) useful in
PT modifying vitamin D receptor activity, in producing anti-VDRE-BP
PT antibodies, in identifying agonists and antagonists of the protein, or
PT in gene therapy
XX
PS Examples; Page 54; 81pp; English.
XX
CC The present invention provides the protein and coding sequences for two
CC vitamin D response element binding proteins (VDRE-BPs) from the marmoset.
CC These confer vitamin D resistance on the primate, a characteristic which
CC2 is associated with high circulating levels of other steroid hormones. The
CC2 sequences provided by the invention can be used to identify treatments
CC2 for osteoporosis, hypercalcaemia, vitamin D intoxication, steroid hormone
CC2 hypersecretion, glucocorticoid mediated disorders, lymphoma and granuloma
CC2 forming diseases. The present sequence is a tryptic peptide derived from
CC2 one of the VDRE-BPs of the invention.
XX
SQ Sequence 7 AA;

Query Match 54.3%; Score 25; DB 22; Length 7;

Best Local Similarity 57.1%; Prod. No. 6.4e-05;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 YDDQESV 7
:|: ::
Db 1 tddkdsr 7

RESULT 4
AAP92282
ID AAP92282 standard; protein; 8 AA.
XX
AC AAP92282;
XX
DT 27-JUN-1994 (first entry)
XX
DE Linkage peptide for subunits I and II of single chain recombinant monellin
DE analogue.
XX
KW Monellin; fusion peptide; sweetener; linkage peptide.
XX
OS Synthetic.
XX
PN WO9815571 A.
XX
PD 29-DEC-1988.
XX
PF 17-JUN-1988; 88WO-US02114.
XX
PR 19-JUN-1987; 87US-0064341.
XX
PA (LUCK) LUCKY BIOTECH CORPORATION.
XX
PI Kim SH, Cho JM;
XX
DR WPI; 1999-238822/03.
XX
PT Sweetener compsn. contg. single chain monellin analogue
PT having the two sub units covalently bonded via peptide link.
XX
PS Claim 14; page 18; 22pp; English.
XX
CC A fusion protein was produced by synthesising 14 oligomers and ligating
CC them to produce a ds DNA mol. encoding the amino acid sequence. The new
CC gene was cloned and used to transform E.coli JM101 to express the
CC protein the linkage peptide, which joins the two subunits of monellin, is
CC one of twelve claimed.
CC The new peptide is a sweetener and can be used in carbonated drinks and
CC various food prods. eg chewing gum, meat prods, instant soups, yoghurt,
CC desserts, and cereal; in animal feeds; in dental hygiene prods. such as
CC toothpaste, mouthwash and pharmaceuticals. It can also be expressed
CC in yeasts or plants to increase sweetness of fruit and vegetables.
CC See also AAP92280-92, and WO8810303-A and WO8815027 A.
XX
SQ Sequence 8 AA;

Query Match 52.2%; Score 24; DB 1; Length 8;
Best Local Similarity 57.1%; Prod. No. 6.4e-05;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 YDDQESVK 8
:|: ::|
Db 1 tddkdsrk 8

RESULT 5
AAR88034
ID AAR88034 standard; Peptide; 8 AA.
XX
AC AAR88034;
XX
DT 10-JUL-1996 (first entry)
XX
DE Linker sequence in monellin single chain 1-alm and. rec.
XX
KW monellin; analogue; Serendipity Berry; 11-score pb.fbp; commutisii;
KW sweetener; food; beverage; single chain form.
XX
OS Synthetic.
XX
PN US5478923 A.
XX
PD 26-DEC-1995.
XX
PF 19-JUN-1987; 87US-0064341.
XX
PR 30-MAR-1990; 90US-0502257.
PR 19-JUN-1987; 87US-0064341.
PR 04-NOV-1987; 87US-0117124.
PR 18-JAN-1990; 90US-0465585.
PR 25-MAR-1994; 94US-0218264.
XX
PA (LUCK) LUCKY CO LTD.
PA (REGC) UNIV CALIFORNIA.
XX
PI Cho JM, Kim S;
XX
DR WPI; 1996-057806/06.
XX
PT New single chain analogues of monellin useful as sweeteners in
PT foods, beverages, pharmaceuticals etc., are stable at high temps.
PT and acidic pH
XX
PS Claim 7; Column 14; 10pp; English.
XX
CC Single chain analogues of the sweet protein, monellin, comprise a
CC peptide homologous to residues 1-46 of the B chain of monellin
CC (AAR88034) linked through the C-terminus to a covalent linker which is in
CC turn linked to the N-terminus of a peptide homologous to residues 1-45
CC of the native A chain (AAR88035) of monellin. Monellin is isolated from
CC "Serendipity Berries" of the West African plant Dioscoreophyllum
CC comminsii. The protein can be used as a substitute for sugar in
CC sweetening foods and beverages. The single chain form of monellin
CC are capable of maintaining their 3-D conformation under conditions which
CC would ordinarily denature the native protein and retain their sweet
CC taste. AAR89037-49 represent covalent linkers used in the analogues.
XX
SQ Sequence 8 AA;

Query Match 52.2%; Score 24; DB 1; Length 8;
Best Local Similarity 57.1%; Prod. No. 6.4e-05;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 YDDQESVK 8
:|: ::|
Db 1 tddkdsrk 8

RESULT 6
AAW32245
ID AAW32245 standard; Protein; 8 AA.
XX
AC AAW32245;
XX
DT 20-APR-1998 (first entry)
XX
DE Linker peptide 4 of a single chain monellin and its
XX
KW Monellin; analogue; thermostable; sweetener; heat tolerant;
KW linker.
XX
OS Synthetic.
XX
PN US5672372 A.

XX
PN [illegible]
XX
PF [illegible]
XX
PR [illegible]
PR [illegible]
PR [illegible]
PR [illegible]
PR [illegible]
PR [illegible]
XX
PA ([illegible]) LUCKY BIOTECH CORP.
PA ([illegible]) UNIV CALIFORNIA.
XX
PI Cho JM, Kim S;
XX
DR WPI; 1997-479455/44.
XX
PT New DNA encoding a single chain form of the sweetener monellin
PT that is stable to heating at acidic pH, useful as additive for
PT foods, beverages, pharmaceuticals, toothpaste etc.
XX
PS Example 5; Column 12; 10pp; English.
XX
CC This linker sequence is used to prepare a recombinant single chain form
CC of the protein sweetener monellin. The recombinant protein has the
CC formula B-C-A where B and A represent peptide portions of the B chain and
CC A chain respectively of the native monellin and C is a linker between the
CC B and A chains. The linker sequence C can be a covalent bond or can
CC consist of a peptide of 1-10 gene encoded amino acids, with sufficient
CC hydrophilicity and length to reside on the outside of the molecule
CC without distortion of the native conformation. Polyethylene glycol
CC oligomers can also be used as linkers. The single chain protein sweetener
CC retains the sweetness properties of the double chain native monellin, but
CC are stable under conditions where the native protein is denatured. This
CC protein sweetener is many times sweeter than sucrose, and can be used as
CC a substitute for sugar in foods and beverages (juices, hot drinks or
CC carbonated beverages), chewing gum, toothpastes, mouthwashes, dental
CC hygiene products and pharmaceuticals. They can also be used to generate
CC specific antibodies for their purification.
XX
SQ Sequence 8 AA;

  Query Match [illegible] Score 24; DB 10; Length 8;
  Best Local Similarity [illegible] Pred. No. [illegible];
  Matches [illegible] Conservative [illegible] Mismatches [illegible] Indels [illegible] Gaps [illegible];

Qy [illegible]
Db [illegible]

RESULT [illegible]
AAW[illegible]
ID AAW[illegible] standard; Protein; 8 AA.
XX
AC AAW[illegible];
XX
DT [illegible] (first entry)
XX
DE Linker sequence 1 of single-chain protein sweetener monellin.
XX
KW [illegible]; Basic protein; sweetener; monellin; heat stability; additive;
KW beverage; pharmaceutical; toothpaste; dental hygiene; antibody;
KW [illegible]; [illegible]; sucrose.
XX
OS [illegible].
XX
PN [illegible]-A.
XX
PD [illegible].
XX
PF [illegible].

XX
PR 30-MAR-1990; 90US-058225?.
PR 19-JUN-1987; 87US-0064441.
PR 19-JUN-1987; 87US-0064443.
PR 04-NOV-1987; 87US-0117124.
PR 18-JAN-1990; 90US-0465585.
PR 02-NOV-1993; 93US-0146326.
PR 06-JUN-1995; 95US-0650545.
XX
PA (LUCK-) LUCKY BIOTECH CORP.
PA (REGC-) UNIV CALIFORNIA.
XX
PI Cho JM, Kim S;
XX
DR WPI; 1997-479455/44.
XX
PT New DNA encoding a single chain form of the sweetener monellin
PT that is stable to heating at acidic pH, useful as additive for
PT foods, beverages, pharmaceuticals, toothpaste etc.
XX
PS Example 5; Column 12; 10pp; English.
XX
CC This linker sequence is used to prepare a recombinant single chain form
CC of the protein sweetener monellin. The recombinant protein has the
CC formula B-C-A where B and A represent peptide portions of the B chain and
CC A chain respectively of the native monellin and C is a linker between the
CC B and A chains. The linker sequence C can be a covalent bond or can
CC consist of a peptide of 1-10 gene encoded amino acids, with sufficient
CC hydrophilicity and length to reside on the outside of the molecule
CC without distortion of the native conformation. Polyethylene glycol
CC oligomers can also be used as linkers. The single chain protein sweetener
CC retains the sweetness properties of the double chain native monellin, but
CC are stable under conditions where the native protein is denatured. This
CC protein sweetener is many times sweeter than sucrose, and can be used as
CC a substitute for sugar in foods and beverages (juices, hot drinks or
CC carbonated beverages), chewing gum, toothpastes, mouthwashes, dental
CC hygiene products and pharmaceuticals. They can also be used to generate
CC specific antibodies for their purification.
XX
SQ Sequence 8 AA;

  Query Match 52.2%; Score 24; DB 10; Length 8;
  Best Local Similarity 37.5%; Pred. No. 6.4e+06;
  Matches 3; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Qy 1 YDDQESVK 8
    |::|: :|
Db 1 yenredik 8

RESULT 8
AAY43571
ID AAY43571 standard; peptide; 6 AA.
XX
AC AAY43571;
XX
DT 26-JAN-2000 (first entry)
XX
DE Scrambled peptide derived from filamentous hemagglutinin.
XX
KW Filamentous hemagglutinin; FHA; Factor X; leukocyte;
KW blood vessel endothelial cell; migration; inflamed tissue;
KW non fimbrial surface associated protein; Bordetella pertussis; CDi;
KW RGD tripeptide; integrin receptor; inflammation; antibiotic therapy;
KW infection; meningitis; septic arthritis; endophthalmitis;
KW autoimmune disease.
XX
OS Synthetic.
OS Bordetella pertussis.
XX
PN US5968512-A.
XX

PD 10-OCT-1995;
XX
PF 06-JUN-1995; 95US-0465966.
XX
PR 30-NOV-1994; 94US-0348353.
PR 04-MAY-1992; 92WO-US03725.
PR 03-MAY-1991; 91US-0695613.
PR 23-MAY-1994; 94US-0247572.
XX
PA (UYRQ) UNIV ROCKEFELLER.
XX
PI Masure HR, Inumaneni E;
XX
DR WPI; 1995-381257/51.
XX
PT Peptides inhibiting the influx of leukocytes into inflamed tissue, useful for dampening inflammation during treatment with antibiotics.
XX
PS Example 9; Column 32; 81pp; English.
XX
CC The present sequence represents a scrambled peptide of filamentous
CC hemagglutinin (FHA). FHA is a non-fimbrial surface associated protein.
CC As there are four regions in FHA with sequence similarity to three
CC regions in Factor X, some antibodies to these four regions cross-react
CC with Factor X, and vice versa. The specification describes FHA peptides
CC which interact with leukocytes or with blood vessel endothelial cells,
CC thereby inhibiting the migration of leukocytes from the blood stream
CC into inflamed tissue. FHA peptides competitively inhibit binding of
CC Factor X and C3bi to leukocytes. Specifically, the peptides inhibit the
CC reaction between the RGD tripeptide of FHA and the integrin receptors of
CC endothelial cells. The FHA peptides (and antibodies to them) dampen
CC inflammation during the course of therapy with antibiotics and are
CC therefore useful in the treatment of infections (e.g. meningitis, septic
CC arthritis, and endophthalmitis), and inflammation arising from autoimmune
CC disease.
XX
SQ Sequence 6 AA;

Query Match 50.0%; Score 23; DB 20; Length 6;
Best Local Similarity 66.7%; Pred. No. 6.4e+05;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 3 DPESVK 8
||:|||
Db 1 DEETVK 6

RESULT 9
AAY23962
ID AAY23962 standard; peptide; 6 AA.
XX
AC AAY23962;
XX
DT 22-SEP-1995 (first entry)
XX
DE Scrambled filamentous haemagglutinin peptide.
XX
KW Adhesion; leukocyte; endothelial cell; bacteria; Bordetella pertussis;
KW ciliated respiratory epithelial cell; filamentous haemagglutinin; FHA;
KW RGD region; cerebral endothelial cell; inflammation; antibiotic therapy;
KW infection; meningitis; septic arthritis; endophthalmitis.
XX
OS Synthetic.
OS Bordetella pertussis.
XX
PN US5444127-A.
XX
PD 22-AUG-1995.
XX
PF 30-NOV-1994; 94US-0348353.
XX
PR 30-NOV-1994; 94US-0348353.
PR 04-MAY-1992; 92WO-US03725.
PR 03-MAY-1991; 91US-0695613.
PR 04-MAY-1992; 92WO-US03725.
PR 23-MAY-1994; 94US-0247572.
XX
PA (UYRQ) UNIV ROCKEFELLER.
XX
PI Masure HR, Inumaneni E;
XX
DR WPI; 1995-443571/47.
XX
PT Peptides inhibiting the adhesion between leukocytes and endothelial
PT cells, useful for treating inflammation
XX
PS Example 9; Column 28; 82pp; English.
XX
CC The specification describes peptides which inhibit adhesion between
CC leukocytes and endothelial cells, and peptides which inhibit adhesion
CC between bacteria (Bordetella pertussis) and ciliated respiratory
CC epithelial cells. The specification also describes an immunogenic
CC composition comprising a polypeptide portion of Bordetella pertussis
CC filamentous haemagglutinin (FHA) containing no RGD region or containing
CC an amino acid sequence altered in the RGD region, where the polypeptide
CC portion elicits antibodies which do not cross react with cerebral
CC endothelial cells. The peptides and methods are useful for reducing
CC inflammation during the course of antibiotic therapy of infectious
CC diseases such as meningitis, septic arthritis, and endophthalmitis.
CC The present sequence represents a scrambled FHA peptide of
CC B. pertussis.
XX
SQ Sequence 6 AA;

Query Match 50.0%; Score 23; DB 20; Length 6;
Best Local Similarity 66.7%; Pred. No. 6.4e+05;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 3 DPESVK 8
||:|||
Db 1 DEETVK 6

RESULT 10
AAY67529
ID AAY67529 standard; peptide; 6 AA.
XX
AC AAY67529;
XX
DT 30-MAY-2000 (first entry)
XX
DE Scrambled FHA peptide used as a control.
XX
KW FHA; blood brain barrier; filamentous hemagglutinin; endothelial cell;
KW complement C3bi; factor X; integrin receptor CR3; leukocyte migration;
KW bacterial adhesion; brain cancer; acquired immune deficiency syndrome;
KW Parkinson's disease; Alzheimer's disease; multiple sclerosis; anti-epileptic;
KW anti-inflammatory; anticancer; antiviral; autoimmunodegeneration
XX
OS Bordetella pertussis.
XX
PN US6015560-A.
XX
PD 18-JAN-2000.
XX
PF 06-JUN-1995; 95US-0465966.
XX
PR 30-NOV-1994; 94US-0348353.
PR 04-MAY-1992; 92WO-US03725.
PR 03-MAY-1991; 91US-0695613.
PR 23-MAY-1994; 94US-0247572.
XX
PA (UYRQ) UNIV ROCKEFELLER.
XX
PI Masure HR, Inumaneni E;

XX
[illegible]
XX
PT [illegible]
[illegible]
XX
PS Example 1; Page 8; 82pp; English.
XX
CC The invention provides a novel pharmaceutical composition for increasing
CC the permeability of the blood brain barrier to a molecule (I). The
CC [illegible] comprises (I) and an antibody to FHA (filamentous
CC hemagglutinin) which binds to endothelial cells in brain blood vessels,
CC to increase permeability. FHA contains polypeptide regions with binding
CC properties similar to those of complement C3bi, factor X and integrin
CC receptor CR3, and some anti-FHA antibodies are competitive inhibitors
CC of Bordetella bacteria, i.e. they reduce leukocyte migration or bacterial
CC adhesion. The compositions are used to improve delivery of (I) to the
CC brain, e.g. where (I) is used to treat brain cancer, acquired immune
CC deficiency syndrome, epilepsy, Parkinson's or Alzheimer's diseases or
CC other neurological diseases. Other antibodies directed against particular
CC regions of FHA are used to treat inflammation (caused by microbial
CC infection or auto immune disease), also to prevent adhesion of Bordetella
CC pertussis to respiratory endothelial cells. The present sequence
CC represents a scrambled FHA peptide used as a control in experiments of
CC leukocyte endothelial cell interactions by FHA derived peptides.
XX
SQ Sequence 9 AA;

Query Match 50.0%; Score 24; DB 22; Length 9;
Best Local Similarity 66.7%; Pred. No. 6.4e+05;
Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 YDDQES 6
|||
Db 3 yddnss 8

RESULT 12
AAB69861
ID AAB69861 standard; peptide; 9 AA.
XX
AC AAB69861;
XX
DT 24-MAR-1995 (first entry)
XX
DE Peptide (232) inhibitor of prolyl endopeptidase.
XX
KW prolyl endopeptidase; inhibitor; pig; kidney.
XX
OS Synthetic.
XX
PN JP06192298-A.
XX
PD 12-JUL-1994.
XX
PF 24-DEC-1992; 92JP-0344855.
XX
PR 24-DEC-1992; 92JP-0344855.
XX
PA (MITK) MITSUI TOATSU CHEM INC.
XX
DR WPI; 1994-260515/32.
XX
PT New peptide - useful for inhibiting prolyl endopeptidase
XX
PS Example; Page 27; 30pp; Japanese.
XX
CC This shows the sequence of a peptide inhibitor of prolyl
CC endopeptidase. The prolyl endopeptidase was isolated and purified
CC from pig kidneys. AAB57582-754 are specific examples of the
CC generic sequence shown in AAB57581.
XX
SQ Sequence 9 AA;

Query Match 47.8%; Score 22; DB 15; Length 9;
Best Local Similarity 57.1%; Pred. No. 6.4e+05;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 2 DDQESVK 8
| : :||
Db 3 ddkanvk 9

RESULT [illegible]
AAY4[illegible]
ID AAY4[illegible] standard; Peptide; 9 AA.
XX
AC AAY4[illegible];
XX
DT [illegible]-DEC-[illegible] (first entry)
XX
DE Immunogenic peptide having a human leukocyte antigen binding motif #1243.
XX
KW Human leukocyte antigen; binding; immunogenic; atypical cells; MHC; HLA;
KW immune response; T cell activation; major histocompatibility complex;
KW cytotoxic T lymphocyte; CTL; tumour rejection; viral infection; cancer;
KW prostate cancer; hepatitis B; hepatitis C; AIDS; renal carcinoma;
KW vaccine; immunisation.
XX
OS Synthetic.
OS Homo sapiens.
XX
PN W[illegible]45[illegible]54-A1.
XX
PD 16-SEP-[illegible].
XX
PF 13-MAR-[illegible]; 98WO-US05039.
XX
PR 14-MAR-[illegible]; 98WO-US05039.
XX
PA (EPIM-) EPIMMUNE INC.
XX
PI Sette A, Fikes RT, Sidney J, Celis E, Grey HM, 2-[illegible]
XX
DR WPI; [illegible]-551211/46.
XX
PT New immunogenic peptides with HLA binding motif, useful in treatment
PT and diagnosis of cancers and viral diseases
XX
PS Claim 1; Page 79; 150pp; English.
XX
CC AAY4[illegible] to AAY48214 represent specifically claimed human class I peptides
CC having a human major histocompatibility complex (MHC) class I (also
CC known as human leukocyte antigen (HLA)) binding motif. The immunogenic
CC peptides can bind to a specific HLA allele (i.e. HLA A subtypes
CC HLA A*0.1, A1, A3.2 or A24.1 or HLA B or C) and induce a cytotoxic T cell
CC response against the antigen from which the peptide is derived.
CC Cytotoxic T lymphocytes (CTLs) which destroy antigen bearing cells are
CC normally induced by an antigen in the form of a peptide fragment bound
CC to a HLA molecule, rather than the intact foreign antigen itself, and
CC are particularly important in tumour rejection and in clearing viral
CC infections. The peptides are therefore useful therapeutically to treat
CC or prevent viral infections and cancers in mammals (especially humans)
CC e.g. prostate cancer, hepatitis B and C, AIDS, and renal carcinoma.
CC They can be administered as vaccines to elicit an immune response in
CC individuals susceptible or otherwise at risk of viral infection or
CC cancer, or used to treat chronic or acute conditions. They are also
CC useful diagnostically, and can be used to induce a cytotoxic T cell
CC response, by contacting a cytotoxic T cell with the peptide e.g. to
CC produce CTLs ex vivo for infusion back into a patient. The
CC polynucleotides encoding the immunogenic peptides are also useful
CC therapeutically and for immunisation as above.
XX
SQ Sequence [illegible] AA;

Query Match 47.8%; Score 22; DB 20; Length [illegible];
Best Local Similarity 50.0%; Pred. No. 6.4e+0[illegible];
Matches 4; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 YIDQPSVK 8
|| | : ||
Db 2 ydlpsvik 9

RESULT 14
AAP81908
ID AAP81908 standard, peptide, 7 AA.
XX
AC AAP81908;
XX
DT 10-OCT-1990 (first entry)
XX
DE Peptide conteg D4K sequence of activation peptide of human trypsinogens.
XX
KW activation peptides of pancreatic zymogens; D4K peptapeptide;
KW pancreatic disease;
XX
OS synthetic.
XX
PN WO8801059-A.
XX
PD 11-FEB-1988.
XX
PF 27-JUL-1987; 87WO-GB00532.
XX
PR 28-JUL-1986; 86GB-0018333.
XX
PA (HERM/) HERMAN TAYLOR J.
PA (SGEO-) ST GEORGES HOSP MED.
PA (BIOS-) BIOSCIENCE INT INC.
XX
PI Hermon Taylor J, Auston BM;
XX
DR WPI; 1988-050042/07.
XX
PT Diagnosis of pancreatic tissue
PT by assaying body fluid for activation peptides of pancreatic
PT zymogens specifically cleaved by proteolysis during activation
XX
PS Disclosure; 7 6pp; English.
XX
CC The Tyr residue was added to allow cross linking of peptide to
CC protein carriers such as bovine serum albumen [illegible] to immunise
CC rabbits.
XX
SQ Sequence 6 AA;

Query Match 45.7%; Score 21; DB 4; Length 6;
Best Local Similarity 60.0%; Pred. No. 2.4e+0[illegible];
Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 YIDQPK 6
||| :
Db 1 ydddd 5

RESULT 15
AAR46052
ID AAR46052 standard, peptide, 7 AA.
XX
AC AAR46052;
XX
DT 26-JUL-1994 (first entry)
XX
DE DR3-restricted T helper cell blocking peptide.
XX
KW Blocker; inhibitor; T cell; T-helper cell; DR3; autoimmune disease;
KW type I diabetes; immunosuppression; Grave's disease;
KW Sjogrens syndrome; myasthenia gravis.
XX
OS Synthetic.
XX
PN WO9402509-A.
XX
PD 03-FEB-1994.
XX

RL 21 [illegible] 93WO NL9201511.

XX

RK 15-OCT-1992, 92EP-0201985.

XX

PA (UYLE-) STICHTSUNIV LEIDEN.

XX

PI Boots JWP, Schuk A, Ottenhoff THM;

XX

DR WPI; 1994-104794/13.

XX

PT New peptides which block self reactive helper T cells
PT restricted to DR4, for treating or preventing autoimmune
PT diseases, e.g. type I diabetes, without causing generalised
PT immunosuppression.

XX

PS Claim 5; Page 15; 64pp; English.

XX

CC Peptides (AA84-92 no57) are specific examples of a peptide
CC I to V, the peptides prevent the activation of self reactive
CC helper cells which lead to the destruction of "self" cells
CC and tissues. The peptides bind to a particular HLA DR molecule
CC but are not recognised by the self reactive T helper cells.

XX

SQ Sequence [illegible] AA;

Query Match 45.7%; Score 21; DB 15; Length 7;
Best Local Similarity 75.0%; Prob. No. 6.4e-05;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YDDQV 4
: |:
Db 4 YDDEV 7

Search completed: August 26, 2002, 06:27:09
CPU time [illegible]

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 26, 2002, 06:23:21 ; Search time 51.79 Seconds
(without alignments)
4.249 Million cell updates/sec

Title: US-09-744-282-3
Perfect score: 46
Sequence: 1 YDDQESVKS 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 463256

Minimum DB seq length: 1
Maximum DB seq length: 9

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/ina/PA_COMB.pep: *
2: /cgn2_6/ptodata/1/ina/PU_COMB.pep: *
3: /cgn2_6/ptodata/1/ina/PA_COMB.pep:*
4: /cgn2_6/ptodata/1/ina/PB_COMB.pep: *
5: /cgn2_6/ptodata/1/ina/5A_COMB.pep: *
6: /cgn2_6/ptodata/1/ina/backfiles1.pep: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	46	100.0	9	1	US-08-619-645-3	Sequence 3, Appli
2	46	100.0	9	2	US-08-624-494-3	Sequence 3, Appli
3	24	52.2	8	1	US-08-483-415-8	Sequence 8, Appli
4	24	52.2	8	6	5487983-6	Patent No. 548798
5	23	50.0	6	2	US-08-348-354-29	Sequence 29, Appl
6	23	50.0	6	2	US-08-465-965-29	Sequence 29, Appl
7	23	50.0	6	3	US-08-465-966-29	Sequence 29, Appl
8	22	47.8	6	1	US-08-200-900A-13	Sequence 13, Appl
9	22	47.8	6	5	PCT US94-00516-13	Sequence 13, Appl
10	21	45.7	7	1	US-08-525-654A-142	Sequence 142, App
11	21	45.7	8	6	5487983-34	Patent No. 548798
12	20	43.5	5	1	US-07-657-769B-20	Sequence 20, Appl
13	20	43.5	6	3	US-08-207-395-75	Sequence 75, Appl
14	20	43.5	7	1	US-07-956-848A-58	Sequence 58, Appl
15	20	43.5	7	1	US-08-471-956-58	Sequence 58, Appl
16	20	43.5	7	4	US-08-665-644A-16	Sequence 16, Appl
17	20	43.5	8	1	US-08-483-415-7	Sequence 7, Appli
18	20	43.5	8	1	US-08-483-415-19	Sequence 19, Appl
19	20	43.5	8	6	5487983-5	Patent No. 548798
20	20	43.5	9	2	US-08-696-944-11	Sequence 11, Appl
21	20	43.5	9	4	US-08-448-489-4	Sequence 4, Appli
22	20	43.5	9	4	US-09-000-041A-1	Sequence 1, Appl
23	20	43.5	9	4	US-08-786-459B-5	Sequence 5, Appli
24	19	41.3	5	1	US-08-704-170-16	Sequence 16, Appl
25	19	41.3	5	5	PCT US94-02641-16	Sequence 16, Appl
26	19	41.3	6	1	US-08-704-170-17	Sequence 17, Appl
27	19	41.3	6	5	PCT US94-02641-17	Sequence 17, Appl
28	19	41.3	7	1	US-08-281-193-9	Sequence 9, Appli
29	19	41.3	7	1	US-08-422-[illegible]-9	Sequence 9, Appl
30	19	41.3	7	2	US-08-745-710-9	Sequence 9, Appli
31	19	41.3	7	2	US-08-[illegible]-9	Sequence 9, Appli
32	19	41.3	7	4	US-09-519-[illegible]-9	Sequence 9, Appli
33	19	41.3	7	4	US-08-665-644A-11	Sequence 11, Appl
34	19	41.3	7	5	PCT US94-03370-7	Sequence 7, Appli
35	19	41.3	7	5	PCT US95-08069-6	Sequence 6, Appli
36	19	41.3	8	1	US-08-014-429-25	Sequence 25, Appl
37	19	41.3	8	1	US-08-483-415-9	Sequence 9, Appli
38	19	41.3	8	2	US-08-467-049-16	Sequence 16, Appl
39	19	41.3	8	3	US-08-835-783-64	Sequence 64, Appl
40	19	41.3	8	4	US-08-967-780-105	Sequence 105, App
41	19	41.3	8	4	US-09-074-898-105	Sequence 105, App
42	19	41.3	8	5	PCT US94-01414-25	Sequence 25, Appl
43	19	41.3	8	6	5487983-7	Patent No. 548798
44	19	41.3	9	1	US-08-704-170-7	Sequence 7, Appl
45	19	41.3	9	2	US-08-657-442-3	Sequence 3, Appl

ALIGNMENTS

RESULT 1
US-08-619-645-3
; Sequence 3, Application US/08619645
; Patent No. 5736507
; GENERAL INFORMATION:
; APPLICANT: Boots, Anna M.H.
; APPLICANT: Verheijden, Gijsbertus F.M.
; TITLE OF INVENTION: [illegible] pept[illegible] bos domiva[illegible]
; TITLE OF INVENTION: [illegible] [illegible] [illegible] [illegible] therapy [illegible]
; TITLE OF INVENTION: disease
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Abbo No. 5736507 Patent Department
; STREET: 1500 Piccard Drive, Suite 206
; CITY: Rockville
; STATE: Maryland
; COUNTRY: U.S.A.
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC DOS/MS DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 EPO
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/619,645
; FILING DATE: 25-MAR-1996
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Gormley, Mary E.
; REGISTRATION NUMBER: 34,409
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 947-4433
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-619-645-3

Query Match 100.0%; Score 46; DB 1; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.2e-05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YDDQESVKS 9
|||||||||
Db 1 YDDQESVKS 9

RESULT 3
US-08-744-282-3
; Sequence 3, Application US/08744282
; Patent No. 5844044
; GENERAL INFORMATION:
; APPLICANT: A.M.H. Boots
; APPLICANT: S.F.M. Verbeijden
; APPLICANT: F.G. Bos
; TITLE OF INVENTION: No. 5844044 cel Peptides derived fr an autoantigen for use
; TITLE OF INVENTION: in Immunotherapy of Autoimmune Diseases
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Akzo No. 5844044cel Patent Department
; STREET: 1300 Piccard Drive, Suite 206
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/744,282
; FILING DATE: 18-APR-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/616,645
; FILING DATE: 25-MAR-1996
; PR APPLICATION DATA:
; APPLICATION NUMBER: EP/EP95/04201
; FILING DATE: 25-OCT-1995
; PR APPLICATION DATA:
; APPLICATION NUMBER: NL 9420[illegible]287
; FILING DATE: 27-OCT-1994
; PRI APPLICATION DATA:
; APPLICATION NUMBER: NL 95[illegible]
; FILING DATE: 07-APR-1995
; ATTORNEY/AGENT INFORMATION:
; TELEPHONE: (301) 258-5200
; TELEFAX: (301) 977-0847
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-744-282-3

Query Match 100.0%; Score 46; DB 2; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.7e-05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YDDQESVKS 9
||||||||| |
Db 1 YDDQESVKS 9

RESULT 4
US-08-744-282-3
; Sequence 3, Application US/08744282
; Patent No. 5844044
; GENERAL INFORMATION:
; APPLICANT: FISCHER, ROBERT
; APPLICANT: KIM, SUNG HO
; APPLICANT: CHO, JOONG M.
; APPLICANT: PENARRUBIA, LOLA
; APPLICANT: GIOVANNONI, JAMES
; APPLICANT: KIM, ROSALIND
; TITLE OF INVENTION: ENHANCEMENTALLY SWEETENED TRANSGENIC PLANT
; TITLE OF INVENTION: PROTEINS
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 2000 PENNSYLVANIA AVENUE, NW
; CITY: WASHINGTON
; STATE: DC
; COUNTRY: USA
; ZIP: 20006-1812
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/484,415
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: MURASHIGE, KATE H.
; REGISTRATION NUMBER: 29,959
; REFERENCE/DOCKET NUMBER: 2000-0501.10
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 887-1500
; TELEFAX: (202) 887-0763
; TELEX: 89-4[illegible] MBSNFOERSWSH
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-484-415-8

Query Match 52.2%; Score 24; DB 1; Length 8;
Best Local Similarity 47.5%; Pred. No. 1.7e-05;
Matches 3; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Qy 1 YDDQESVK 8
|:::|.:|
Db 1 YENREDIK 8

RESULT 4
5487983-6
;Patent No. 5487983
; APPLICANT: KIM,SUNG-HOU; CHO,JOONG M.
; TITLE OF INVENTION: EXPRESSION SYSTEMS FOR MAKING
;SINGLE-CHAIN MONELLIN ANALOGS
; NUMBER OF SEQUENCES: 36
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/148,526
; FILING DATE: 02-NOV-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 502,257
; FILING DATE: 30-MAY-1990
; APPLICATION NUMBER: 465,585
; FILING DATE: 18-JAN-1990
; APPLICATION NUMBER: 117,124
; FILING DATE: 04-NOV-1987
; APPLICATION NUMBER: 64,341
; FILING DATE: 19-JUN-1987
; APPLICATION NUMBER: 64,343
; FILING DATE: 19-JUN-19987
;SEQ ID NO:6:
; LENGTH: 8
5487983-6

Query Match 52.2%; Score 24; DB 6; Length 8;
Best Local Similarity 47.5%; Pred. No. 1.7e-05;
Matches 3; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

```
Qy       1 VEVLESVK 8
           ::::::::
Db       1 VENREDIK 8


RESULT   6
US-08-348-354-19
; Sequence 19, Application US/08348354
; Patent No. 5982217
;  GENERAL INFORMATION:
;    APPLICANT:  Thomanen, Elaine
;    APPLICANT:  Masure, Robert
;    TITLE OF INVENTION:  Antibody Recognizing Endothelial Cell
;    TITLE OF INVENTION:  Ligand for Leukocyte CR3
;    NUMBER OF SEQUENCES:  38
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Klauber & Jackson
;      STREET:  411 Hackensack Avenue
;      CITY:  Hackensack
;      STATE:  New Jersey
;      COUNTRY:  U.S.A.
;      ZIP:  07601
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  08/348,354
;      FILING DATE:  09-NOV-1994
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Jackson, David
;      REGISTRATION NUMBER:  26,742
;      REFERENCE/DOCKET NUMBER:  600-1 097CIP1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  201 487-5800
;      TELEFAX:  201 343 1684
;  INFORMATION FOR SEQ ID NO:  19:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  6 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-348-354-19


  Query Match             50.0%;  Score 24;  DB 2;  Length 6;
  Best Local Similarity   66.7%;  Pred. No. 1.7e+05;
  Matches    4;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3 DQESVK 8
           ||:|||
Db       1 DEEIVK 6


RESULT   7
US-08-465-966-19
; Sequence 19, Application US/08465966
; Patent No. 6015560
;  GENERAL INFORMATION:
;    APPLICANT:  Thomanen, Elaine
;    APPLICANT:  Masure, Robert
;    TITLE OF INVENTION:  Antibody Recognizing Endothelial Cell
;    TITLE OF INVENTION:  Ligand for Leukocyte CR3
;    NUMBER OF SEQUENCES:  38
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Klauber & Jackson
;      STREET:  411 Hackensack Avenue
;      CITY:  Hackensack
;      STATE:  New Jersey
;      COUNTRY:  U.S.A.
;      ZIP:  07601
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  08/465,965
;      FILING DATE:  06-JUN-1995
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/348,354
;      FILING DATE:  30-NOV-1994
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/247,572
;      FILING DATE:  23-MAY-1994
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US92/03725
;      FILING DATE:  04-MAY-1992
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/695,613
;      FILING DATE:  03-MAY-1991
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Jackson, David
;      REGISTRATION NUMBER:  26,742
;      REFERENCE/DOCKET NUMBER:  600-1 097CIPDIV2
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  201 487-5800
;      TELEFAX:  201 343 1684
;  INFORMATION FOR SEQ ID NO:  19:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  6 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-465-965-19


  Query Match             50.0%;  Score 24;  DB 2;  Length 6;
  Best Local Similarity   66.7%;  Pred. No. 1.7e+05;
  Matches    4;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3 DQESVK 8
           ||:|||
Db       1 DEEIVK 6


RESULT   7
US-08-465-966-19
; Sequence 19, Application US/08465966
; Patent No. 6015560
;  GENERAL INFORMATION:
;    APPLICANT:  Thomanen, Elaine
;    APPLICANT:  Masure, Robert
;    TITLE OF INVENTION:  Antibody Recognizing Endothelial Cell
;    TITLE OF INVENTION:  Ligand for Leukocyte CR3
;    NUMBER OF SEQUENCES:  38
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Klauber & Jackson
;      STREET:  411 Hackensack Avenue
;      CITY:  Hackensack
;      STATE:  New Jersey
;      COUNTRY:  U.S.A.
;      ZIP:  07601
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
```

```
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  Patentin Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/465,966
;       FILING DATE:  06-JUN-1995
;       CLASSIFICATION:  424
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  08/348,354
;       FILING DATE:  30-NOV-1994
;       CLASSIFICATION:  424
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  08/247,572
;       FILING DATE:  23-MAY-1994
;       CLASSIFICATION:  424
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  PCT/US92/10425
;       FILING DATE:  04-MAY-1992
;       CLASSIFICATION:  424
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  07/695,614
;       FILING DATE:  03-MAY-1991
;       CLASSIFICATION:  424
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Jackson, David
;       REGISTRATION NUMBER:  26,742
;       REFERENCE/DOCKET NUMBER:  690-1 097CIP1.IV2
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  201-487-5800
;       TELEFAX:  201 343 1684
;   INFORMATION FOR SEQ ID NO:  19:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  5 amino acids
;       TYPE:  amino acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  peptide
US-08-4[illegible]-A-19

  Query Match                 50.0%;  Score 23;  DB 5;  Length 6;
  Best Local Similarity   66.7%;  Pred. No. 1.7e+05;
  Matches    4;  Conservative    2;  Mismatches    0;  Indels    1;  Gaps    0;

Qy       2 [illegible]
           :|||:
Db       2 DEEIVK 6


RESULT  [illegible]
US-[illegible]-A-[illegible]
; Sequence 19, Application US/08200900A
; Patent No. 5665566
;  GENERAL INFORMATION:
;     APPLICANT:
;     TITLE OF INVENTION:  CLONING OF ENTEROKINASE AND METHOD OF USE
;     NUMBER OF SEQUENCES:  88
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Genetics Institute, Inc.  Legal Affairs
;       STREET:  87 CambridgePark Drive
;       CITY:  Cambridge
;       STATE:  MA
;       COUNTRY:  USA
;       ZIP:  02140
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  Patentin Release #1.0, Version #1.25 (EPO)
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/200,900A
;       FILING DATE:  25-FEB-1994
;       CLASSIFICATION:  435
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  McInerney, Maureen C.
;       REGISTRATION NUMBER:  41,544
;       REFERENCE/DOCKET NUMBER:  GI 5201-PWC
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (617) 876-1170 X8774
;       TELEFAX:  (617) 876-5851
;   INFORMATION FOR SEQ ID NO:  10:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  6 amino acids
;       TYPE:  amino acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  protein
US-08-200,900A-10


  Query Match                 47.8%;  Score 22;  DB 1;  Length 6;
  Best Local Similarity   60.0%;  Pred. No. 1.7e+05;
  Matches    3;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 YDDQE 5
           :|||:
Db       1 FDDQQ 5


RESULT  9
PCT-US94-00616-10
; Sequence 10, Application PC/TUS9400616
;  GENERAL INFORMATION:
;     APPLICANT:
;     TITLE OF INVENTION:  CLONING OF ENTEROKINASE AND METHOD OF USE
;     NUMBER OF SEQUENCES:  33
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  Patentin Release #1.0, Version #1.25 (EPO)
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  PCT/US94/00616
;       FILING DATE:
;       CLASSIFICATION:
;   INFORMATION FOR SEQ ID NO:  10:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  6 amino acids
;       TYPE:  amino acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  protein
PCT-US94-00616-10


  Query Match                 47.8%;  Score 22;  DB 5;  Length 6;
  Best Local Similarity   60.0%;  Pred. No. 1.7e+05;
  Matches    3;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 YDDQE 5
           :|||:
Db       1 FDDQQ 5


RESULT  10
US-08-525,654A-142
; Sequence 142, Application US/08525654A
; Patent No. 5736456
;  GENERAL INFORMATION:
;     APPLICANT:  SANO, KOICHIRO
;     APPLICANT:  KUMAZAWA, YOSHIYUKI
;     APPLICANT:  YASEUDA, HISASHI
;     APPLICANT:  SEGURO, KATSUYA
;     APPLICANT:  MORIKI, MASAO
;     TITLE OF INVENTION:  TRANSGLUTAMINASE ORIGINATED FROM
;     TITLE OF INVENTION:  CHASSOSTREA SIGAS
;     NUMBER OF SEQUENCES:  150
```

```
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT
;       ADDRESSEE:  P.C.
;       STREET:  1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
;       CITY:  ARLINGTON
;       STATE:  VA
;       COUNTRY:  USA
;       ZIP:  22202
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC DOS/MS DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.30
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/525,654A
;       FILING DATE:  28-SEP-1995
;       CLASSIFICATION:  435
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  JP 6/8283
;       FILING DATE:  28-JAN-1994
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  JP 7/7979
;       FILING DATE:  13-JAN-1995
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  OBLON, NORMAN F.
;       REGISTRATION NUMBER:  24,618
;       REFERENCE/DOCKET NUMBER:  HC760 0 PCT
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  703-413-3000
;       TELEFAX:  703-413-2220
;   INFORMATION FOR SEQ ID NO:  142:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  7 amino acids
;       TYPE:  amino acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  peptide
;     ORIGINAL SOURCE:
;       ORGANISM:  Crassostrea gigas
US-08-525,654A-142


  Query Match             45.7%;  Score 21;  DB 1;  Length 7;
  Best Local Similarity   42.9%;  Pred. No. 1.2e+05;
  Matches    3;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 YDDQESV 7
           ::: : :
Db       1 YEENESM 7


RESULT  11
5487983-34
;Patent No. 5487983
;     APPLICANT: KIM,SUNG-HOU; CHO,JOONG M.
;     TITLE OF INVENTION: EXPRESSION SYSTEMS FOR MAKING
;SINGLE CHAIN M NEILLIN ANALOGS
;     NUMBER OF SEQUENCES: 46
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/146,326
;       FILING DATE: 02-NOV-1993
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER: 502,257
;       FILING DATE: 30-MAY-1990
;       APPLICATION NUMBER: 465,585
;       FILING DATE: 18-JAN-1990
;       APPLICATION NUMBER: 117,121
;       FILING DATE: 04-NOV-1987
;       APPLICATION NUMBER: 64,441
;       FILING DATE: 19-JUN-1987
;       APPLICATION NUMBER: 64,443
;       FILING DATE: 19-JUN-1987
;SEQ ID NO 144:
;       LENGTH: 8
5487983-34


  Query Match             45.7%;  Score 21;  DB 1;  Length 8;
  Best Local Similarity   42.9%;  Pred. No. 1.2e+05;
  Matches    3;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       2 DDQESVK 8
           ::: : :
Db       2 DNREDIK 8


RESULT  12
US-07-657-769B-20
; Sequence 20, Application US/07657769B
; Patent No. 5250766
;  GENERAL INFORMATION:
;    APPLICANT:  GOEBELIN, SHAUN K.
;    TITLE OF INVENTION:  RECOMBINANT RECEPTOR AND RELATED
;    TITLE OF INVENTION:  PHARMACEUTICALS
;    NUMBER OF SEQUENCES:  69
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  IRELL & MANELLA
;      STREET:  545 MIDDLEFIELD ROAD, SUITE 200
;      CITY:  MENLO PARK
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94025
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/657,769B
;      FILING DATE:  19910219
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  MURASHIGE, KATE H.
;      REGISTRATION NUMBER:  29,959
;      REFERENCE/DOCKET NUMBER:  2000-0502.0
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415 327 7250
;      TELEFAX:  415-327 2951
;      TELEX:  706141
;  INFORMATION FOR SEQ ID NO:  20:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  5 amino acids
;      TYPE:  AMINO ACID
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-07-657-769B-20


  Query Match             43.5%;  Score 20;  DB 1;  Length 5;
  Best Local Similarity   60.0%;  Pred. No. 1.7e+05;
  Matches    3;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 YDDQE 5
           :::||
Db       1 WEDQE 5


RESULT  13
US-08-297,395-75
; Sequence 75, Application US/08297395A
; Patent No. 6039947
; GENERAL INFORMATION:
;  APPLICANT: Howard L. Weiner
;  APPLICANT: David A. Hafler
;  TITLE OF INVENTION: PEPTIDES DERIVED FROM MYELIN PROTEINS
;  TITLE OF INVENTION: EPITOPES OF MYELIN BASIC PROTEIN
```

; FILE REFERENCE: [illegible]
; PRESENT APPLICATION NUMBER: US/08/[illegible]
; PRESENT FILING DATE: 1994-08-11
; EARLIER APPLICATION NUMBER: 08/059,809
; EARLIER FILING DATE: 1993-05-06
; EARLIER APPLICATION NUMBER: 07/512,259
; EARLIER FILING DATE: 1990-04-20
; EARLIER APPLICATION NUMBER: PCT/US88/02193
; EARLIER FILING DATE: 1988-06-24
; EARLIER APPLICATION NUMBER: 07/065,734
; EARLIER FILING DATE: 1987-06-24
; NUMBER OF SEQ_ID NOS: 84
; SOFTWARE: PatentIn for Windows Version [illegible]
; SEQ_ID NO [illegible]
; LENGTH: [illegible]
; TYPE: PRT
; ORGANISM: Homo sapiens
US-[illegible]

Query Match 44.5%; Score 20; DB 3; Length [illegible];
Best Local Similarity 75.0%; Pred. No. 1.7e+05;
Matches [illegible]; Conservative [illegible]; Mismatches [illegible]; Indels [illegible]; Gaps [illegible];

Qy [illegible]
Db [illegible]

RESULT [illegible]
US-[illegible]
; Sequence [illegible], Application US/[illegible]
; Patent No. [illegible]
; GENERAL INFORMATION:
; APPLICANT: [illegible]
; APPLICANT: [illegible]
; APPLICANT: [illegible]
; APPLICANT: [illegible]
; TITLE OF INVENTION: [illegible]
; NUMBER OF SEQUENCES: [illegible]
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: [illegible]
; STREET: [illegible]
; CITY: [illegible]
; STATE: [illegible]
; COUNTRY: [illegible]
; ZIP: [illegible]
; COMPUTER READABLE FORM:
; MEDIUM TYPE: [illegible]
; COMPUTER: [illegible]
; OPERATING SYSTEM: [illegible]
; SOFTWARE: [illegible]
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: [illegible]
; FILING DATE: [illegible]
; CLASSIFICATION: [illegible]
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: [illegible]
; FILING DATE: [illegible]
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: [illegible]
; FILING DATE: [illegible]
; ATTORNEY/AGENT INFORMATION:
; NAME: [illegible]
; REGISTRATION NUMBER: [illegible]
; REFERENCE/DOCKET NUMBER: [illegible]
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: [illegible]
; TELEFAX: [illegible]
; TELEX: [illegible]
; INFORMATION FOR SEQ ID NO: [illegible]:
; SEQUENCE CHARACTERISTICS:
; LENGTH: [illegible]
; TYPE: [illegible]
; STRANDEDNESS: [illegible]
; TOPOLOGY: [illegible]
; MOLECULE TYPE: [illegible]
US-[illegible]

Query Match 44.5%; Score 20; DB 1; Length 7;
Best Local Similarity 75.0%; Pred. No. 1.7e+05;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YDDG 4
||| :
Db 4 YDDK 7

RESULT 15
US-08-471-956-58
; Sequence 58, Application US/08471956
; Patent No. 5798336
; GENERAL INFORMATION:
; APPLICANT: Travis, James
; APPLICANT: Shafer, William M.
; APPLICANT: Bangalore, Neelesh
; APPLICANT: Pohl, Jan
; TITLE OF INVENTION: Antimicrobial Peptides
; NUMBER OF SEQUENCES: 59
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Greenlee and Winner, P.C.
; STREET: 5370 Manhattan Circle, Suite 201
; CITY: Boulder
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/471,956
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/956,848
; FILING DATE: 02-OCT-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/541,635
; FILING DATE: 21-JUN-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Ferber, Donna M.
; REGISTRATION NUMBER: 33,878
; REFERENCE/DOCKET NUMBER: 3-900
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 499-8080
; TELEFAX: (303) 499-8089
; TELEX: 49617824
; INFORMATION FOR SEQ ID NO: 58:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-471-956-58

Query Match 44.5%; Score 20; DB 1; Length 7;
Best Local Similarity 75.0%; Pred. No. 1.7e+05;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YDDG 4
||| :
Db 4 YDDK 7

Search completed: August 26, 2002, 06:23:22
Job time: 753 sec

GeneCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

SM protein - protein search, using sw model

Run on: August 26, 2002, 06:24:49 ; Search time 7s.0 seconds
(without alignments)
11.702 Million cell updates/sec

Title: US-09-744-282-4
Perfect score: 43
Sequence: 1 ESKIASNTQ 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 28138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 788

Minimum DB seq length: 0
Maximum DB seq length: 9

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_71:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	17	39.5	8	2	149404	prealbumin - weste
2	16	34.9	8	2	A32523	peptidyl-dipeptida
3	15	34.9	9	2	PI0299	Ig heavy chain CRD
4	14	32.6	4	2	A40145	branched-chain-ami
5	14	32.6	5	2	PT0686	T-cell receptor be
6	14	32.6	5	2	PT0659	T-cell receptor be
7	14	32.6	5	2	PI0695	T-cell receptor be
8	14	32.6	7	2	I50210	gene 2 rel protein
9	14	32.6	8	2	PT0639	T-cell receptor be
10	14	32.6	9	2	A45199	L hyosophorin - Ja
11	14	32.6	9	2	PH0902	T-cell receptor be
12	13	30.2	4	2	PT0534	T-cell receptor be
13	13	30.2	4	2	PT0566	T-cell receptor be
14	13	30.2	5	2	PI0625	T-cell receptor be
15	13	30.2	5	2	PT0577	T-cell receptor be
16	13	30.2	6	2	PT0662	T-cell receptor be
17	13	30.2	6	2	PT0723	T-cell receptor be
18	13	30.2	6	2	S71349	beta-crystallin B2
19	13	30.2	7	1	XEYDGD	galactose oxidase
20	12	27.9	5	2	A37114	hypoxanthine phosp
21	12	27.9	6	2	A61411	amoebin - rat
22	12	27.9	7	2	B39127	phosphotransferase
23	12	27.9	8	2	S45651	probable Na+-trans
24	12	27.9	9	2	A11497	transaldolase (EC
25	12	27.9	9	2	PC7073	uniprimol - cytochr
26	12	27.9	9	2	PH0921	T-cell receptor be
27	12	27.9	9	4	I73804	hypothetical E2 pr
28	11	25.6	5	2	PI0561	T-cell receptor be
29	11	25.6	5	2	PT0707	T-cell receptor be
30	11	25.6	5	2	PI0596	T-cell receptor be
31	11	25.6	6	2	I48126	alpha tubulin - ch
32	11	25.6	6	2	PT0514	T-cell receptor be
33	11	25.6	6	2	PI0619	T-cell receptor be
34	11	25.6	7	2	S09027	carboxylesterase (
35	11	25.6	7	2	PI0671	T-cell receptor be
36	11	25.6	8	2	S43971	tumor associated a
37	11	25.6	8	2	S43972	tumor associated a
38	11	25.6	8	2	S21275	cellulase (EC ...
39	11	25.6	8	2	A61467	prealbumin - Adeli
40	11	25.6	9	1	YEPG	thymic factor - pi
41	11	25.6	9	2	A60957	thymocyte growth p
42	11	25.6	9	2	A61386	macrophage inhibit
43	11	25.6	9	2	I52974	seminal vesicle pr
44	11	25.6	9	2	S56004	alkali 1,4-beta-xl
45	11	25.6	9	2	Q85802	hypothetical prote

ALIGNMENTS

RESULT 1
149404
prealbumin - western wild mouse (fragment)
C;Species: Mus spretus (western wild mouse)
C;Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 09-Nov-1999
C;Accession: 149404
R;K., M.S.; Wanst, X.; Hart-G, J.B.; Bulo, M.P.; Takahashi, N.; Mardon, G.; Nadeau,
Mamm. Genome 5, 449-455, 1994
A;Title: Genetic mapping of 40 cDNA clones on the mouse genome by IVF
A;Reference number: 148944; MUID:94345082
A;Accession: 149404
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-8 <RES>
A;Cross-references: EMBL:X59385; NID:g46788; PIDN:CAA41904.1; PID:g42827

Query Match 39.5%; Score 17; DB 2; Length 8;
Best Local Similarity 50.0%; Pred. No. 2.8e+07;
Matches 3; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 4 IASNTQ 9
: | | |
Db 2 VVSNPQ 7

RESULT 2
A32523
peptidyl-dipeptidase A (EC 3.4.15.1) - bovine (fragment)
N;Alternate names: angiotensin I-converting enzyme; peptidyl-dipeptidase I
C;Species: Bos primigenius taurus (cattle)
C;Date: 18-Oct 1989 #sequence_revision 18-Oct-1989 #text_change 11-May-2000
C;Accession: A32523
R;Harris, R.B.
Adv. Exp. Med. Biol. 198, 513-521, 1986
A;Title: Isolation and sequencing of an active-site peptide from angiotensin I-co
A;Reference number: A32523; MUID:87123961
A;Accession: A32523
A;Molecule type: protein
A;Residues: 1-8 <HAR>
C;Superfamily: mammalian peptidyl-dipeptidase A
C;Keywords: alternative splicing; blood pressure control; peptidyl-dipeptidase;

Query Match 34.9%; Score 15; DB 2; Length 8;
Best Local Similarity 28.6%; Pred. No. 1.8e+08;
Matches 2; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ESKIASN 7
: : : : :
Db 1 FIELADS 7

RESULT 3
PT0299
Ig heavy chain CHO3 region (clone 5-103B) - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 30-Sep-1993 #Sequence_revision 30-Sep-1993 #text_change 14-Apr-1995
C;Accession: PT0299
R;Yamada, M.; Wasserman, R.; Reichard, B.A.; Shane, S.; Caton, A.J.; Rovera, G.
J. Exp. Med. 173, 395-407, 1991
A;Title: Preferential utilization of specific immunoglobulin heavy chain diversity and j
A;Reference number: PT0222; MUID:91108437
A;Accession: PT0299
A;Molecule type: DNA
A;Residues: 1-9 <YAM>
A;Experimental source: B lymphocyte
C;Keywords: heterotetramer; immunoglobulin

Query Match 44.9%; Score 15; DB 2; Length 9;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 SNI 8
|||
Db 7 SNI 9

RESULT 4
A40135
branched-chain-amino-acid transaminase (EC 2.6.1.42), mitochondrial - rat (fragment)
N;Alternate names: branched-chain amino-acid aminotransferase, mitochondrial
C;Species: Rattus norvegicus (Norway rat)
C;Date: 18-May-1992 #Sequence_revision 18-May-1992 #text_change 09-Sep-1993
C;Accession: A40135
R;Gibson, S.M.; Wallin, R.; Ball, L.K.
Submitted to the Protein Sequence Database, March 1992
A;Reference number: A40135
A;Accession: A40135
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-4 <BOI>
C;Keywords: aminotransferase; mitochondrion

Query Match 42.6%; Score 14; DB 2; Length 4;
Best Local Similarity 50.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 4 ASN 7
:||:
Db 1 VSSN 4

RESULT 5
PT0686
T-cell receptor beta chain V-D-J region (120-1K and 154-2D) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 17-Jul-1992 #Sequence_revision 17-Jul-1992 #text_change 05-Sep-1997
C;Accession: PT0615; PT0686
R;Feeney, A.J.
J. Exp. Med. 174, 115-124, 1991
A;Title: Junctional sequences of total T cell receptor beta chains have few N regions.
A;Reference number: PT0509; MUID:91277601
A;Accession: PT0615
A;Status: translation not shown
A;Molecule type: mRNA
A;Residues: 1-5 <FEE>
A;Experimental source: newborn thymus; strain BALB/c; clone 120-1k
A;Accession: PT0686
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-5 <FEE>
A;Experimental source: day 18 fetal thymus; strain BALB/c; clone 154-2b
C;Keywords: T-cell receptor

Query Match 42.6%; Score 14; DB 2; Length 5;
Best Local Similarity 75.0%; Pred. No. 2.8e+05;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 5 ASNI 8
|||:|
Db 1 ASSI 4

RESULT 6
PT0659
T-cell receptor beta chain V-D-J region (12.1-8K) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 17-Jul-1992 #Sequence_revision 17-Jul-1992 #text_change 06-May-1997
C;Accession: PT0659
R;Feeney, A.J.
J. Exp. Med. 174, 115-124, 1991
A;Title: Junctional sequences of total T cell receptor beta chains have few N regions
A;Reference number: PT0509; MUID:91277601
A;Accession: PT0659
A;Status: translation not shown
A;Molecule type: mRNA
A;Residues: 1-5 <FEE>
A;Experimental source: day 4 postnatal thymus, strain BALB/c
C;Keywords: T-cell receptor

Query Match 42.6%; Score 14; DB 2; Length 5;
Best Local Similarity 60.0%; Pred. No. 2.8e+05;
Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 5 ASNTQ 9
|||:|
Db 1 ASSAQ 5

RESULT 7
PT0695
T-cell receptor beta chain V-D-J region (135-1D) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 17-Jul-1992 #Sequence_revision 17-Jul-1992 #text_change 06-May-1997
C;Accession: PT0695
R;Feeney, A.J.
J. Exp. Med. 174, 115-124, 1991
A;Title: Junctional sequences of total T cell receptor beta chains have few N regions
A;Reference number: PT0509; MUID:91277601
A;Accession: PT0695
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-5 <FEE>
A;Experimental source: newborn thymus, strain BALB/c
C;Keywords: T-cell receptor

Query Match 42.6%; Score 14; DB 2; Length 5;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 ASN 7
|||
Db 1 ASN 3

RESULT 8
I50210
gene c-rel protein - chicken (fragment)
C;Species: Gallus gallus (chicken)
C;Date: 13-Sep-1996 #Sequence_revision 13-Sep-1996 #text_change 28-Feb-1997
C;Accession: I50210
R;Kabrun, N.; Bumstead, N.; Hayman, M.J.; Enrietto, P.J.

Mol. Cell. Biol. 10, 4788-4794, 1990
A;Title: Characterization of a novel promoter insertion in the c-rel locus.
A;Reference number: I50210; MUID:90455995
A;Accession: I5[illegible]
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-7 <KAR>
A;Cross-references: GB:M55577; NID:g555438; PID:aa11661
C;Genetics:
A;Gene: c-rel

Query Match 32.6%; Score 14; DB 2; Length 7;
Best Local Similarity 75.0%; Pred. No. 2.8e+05;
Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 5 ASNT 8
|:||
Db 4 ADNI 7

RESULT 9
PT0639
T-cell receptor beta chain V-D-J region (111-1AA) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 17-Jul-1992 #sequence_revision 17 Jul 1992 #text_change [illegible] May 1997
C;Accession: PT0639
R;Feeney, A.J.
J. Exp. Med. 174, 115-124, 1991
A;Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.
A;Reference number: PT0509; MUID:91277601
A;Accession: PT0639
A;Status: translation not shown
A;Molecule type: mRNA
A;Residues: 1-8 <FPE>
A;Experimental source: newborn thymus, strain BALB/c
C;Keywords: T-cell receptor

Query Match 32.6%; Score 14; DB 2; Length 8;
Best Local Similarity 75.0%; Pred. No. 2.8e+05;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 5 ASNT 8
|:||
Db 1 ASSI 4

RESULT 10
A45199
L hyosophorin - Java medaka (fragment)
C;Species: Oryzias melastigmus (Java medaka)
C;Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 27-May-1999
C;Accession: A45199
R;Iacnichi, F.; Seko, A.; Kitajima, K.; Inoue, S.; Iwamatsu, T.; Ebe, Y.H.; Morris, H.R.
J. Biol. Chem. 268, 2358-2362, 1993
A;Title: Structural studies of a novel type of tetraantennary sialoglycan unit in a carb
A;Reference number: A45199; MUID:93155046
A;Accession: A45199
A;Molecule type: protein
A;Residues: 1-9 <TAG>
C;Comment: This sialoglycoprotein is 90% carbohydrate by weight. This low molecular weig
C;Keywords: glycoprotein
F;5/Binding site: carbohydrate (Asn) (covalent) #status experimental

Query Match 32.6%; Score 14; DB 2; Length 9;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 ASN 7
|||
Db 3 ASN 5

RESULT 11
PH0902
T-cell receptor beta chain V-D-J reg. 5 (Syr-3'-e-2.8G-4) - rat (fragment)
C;Species: Rattus norvegicus (Norway rat)
C;Date: 09-Oct-1992 #sequence_revision 09-Oct-1992 #text_change 06-May-1997
C;Accession: PH0902
R;Gold, D.P.; Ottner, H.; Sun, D.; Wiley, S.; Vandenbark, A.A.; Wilson, L.B.
J. Exp. Med. 174, 1467-1476, 1991
A;Title: Analysis of T cell receptor beta chains in Lewis rats with experimental aller
A;Reference number: PH0891; MUID:92078017
A;Accession: PH0902
A;Molecule type: mRNA
A;Residues: 1-9 <GOL>
A;Experimental source: myelin basic protein-immunized T cell
C;Keywords: T-cell receptor

Query Match 32.6%; Score 14; DB 2; Length 9;
Best Local Similarity 75.0%; Pred. No. 2.8e+05;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 5 ASNT 8
|:||
Db 2 ASST 5

RESULT 12
PT0534
T-cell receptor beta chain V-D-J region (12F-2A1) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
C;Accession: PT0534
R;Feeney, A.J.
J. Exp. Med. 174, 115-124, 1991
A;Title: Junctional sequences of fetal T cell receptor beta chains have few N regions
A;Reference number: PT0509; MUID:91277601
A;Accession: PT0534
A;Status: translation not shown
A;Molecule type: mRNA
A;Residues: 1-4 <FEE>
A;Experimental source: adult thymus, strain BALB/c
C;Keywords: T-cell receptor

Query Match 30.2%; Score 13; DB 2; Length 4;
Best Local Similarity 75.0%; Pred. No. 2.8e+05;
Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 5 ASNT 8
|| |
Db 1 ASKT 4

RESULT 13
PT0566
T-cell receptor beta chain V-D-J region (14F-2H) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
C;Accession: PT0566
R;Feeney, A.J.
J. Exp. Med. 174, 115-124, 1991
A;Title: Junctional sequences of fetal T cell receptor beta chains have few N regions
A;Reference number: PT0509; MUID:91277601
A;Accession: PT0566
A;Status: translation not shown
A;Molecule type: mRNA
A;Residues: 1-4 <FEE>
A;Experimental source: day 19 fetal thymus, strain BALB/c
C;Keywords: T-cell receptor

Query Match 30.2%; Score 14; DB 2; Length 4;
Best Local Similarity 75.0%; Pred. No. 2.8e+05;
Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy ASN1 6
|| |
Db ASI1 4

RESULT 14
P10625
T cell receptor beta chain V to J region (12c 2S) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 27-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
C;Accession: P10625
R;Feeney, A.J.
J. Exp. Med. 174, 115-124, 1991
A;Title: Functional sequences of fetal T cell receptor beta chains have few N regions.
A;Reference number: P10509; MUID:91277601
A;Accession: P10625
A;Status: translation not shown
A;Molecule type: mRNA
A;Residues: 1-5 <FEE>
A;Experimental source: newborn thymus, strain BALB/c
C;Keywords: T cell receptor

Query Match 30.2%; Score 14; DB 2; Length 5;
Best Local Similarity 60.0%; Pred. No. 2.8e+05;
Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy ASN1Q 6
|| : |
Db ASSL_ 5

RESULT 15
P10577
T cell receptor beta chain V D J region (141 186) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 27-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
C;Accession: P10577; P10574
R;Feeney, A.J.
J. Exp. Med. 174, 115-124, 1991
A;Title: Functional sequences of fetal T cell receptor beta chains have few N regions.
A;Reference number: P10509; MUID:91277601
A;Accession: P10577
A;Status: translation not shown
A;Molecule type: mRNA
A;Residues: 1-5 <FEE>
A;Experimental source: day 19 fetal thymus, strain BALB/c, clone 141 186
A;Accession: P10574
A;Status: translation not shown
A;Molecule type: mRNA
A;Residues: 1-5 <FEE>
A;Experimental source: day 19 fetal thymus, strain BALB/c, clone 111 b_
C;Keywords: T cell receptor

Query Match 30.2%; Score 14; DB 2; Length 5;
Best Local Similarity 60.0%; Pred. No. 2.8e+05;
Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY ASN1_ 6
|| :
Db ASSK_ 5

Search completed: August 26, 2002, 06:24:50
Job time: 95 secs

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

SW protein protein search, using sw model

Run on: August 26, 2002, 06:40:44 ; Search time 44.94 Seconds
(without alignments)
7.754 Million cell updates/sec

Title: US-09-744-282-4
Perfect score: 43
Sequence: 1 FSKIASNTQ 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0
Maximum DB seq length: 9

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	13	30.2	7	1	IGAO_DACDE	P06294 Dactylium d
2	12	27.9	9	1	IAL3_PICJA	P17441 Pichia jadi
3	11	25.6	6	1	DN96_CLOPA	P81561 Clostridium
4	11	25.6	8	1	CAD1_ENTFA	P13278 Enterococcu
5	11	25.6	9	1	IHYF_PIG	P16275 Sus scrofa
6	10	23.3	4	1	FYF1_ANTP	P82776 Acyl phosph
7	10	23.3	5	1	PSK_DAUCA	P69241 Daucus caro
8	10	23.3	5	1	IP2S_CANFA	P51724 Canis famil
9	10	23.3	6	1	LGK1_LOCMI	P11401 Locusta mig
10	10	23.3	7	1	CYA_ENTFA	P13282 Enterococcu
11	10	23.3	7	1	UCP4_MAIZE	P80580 Zea mays (m
12	10	23.3	8	1	AKHG_GRYBI	P14096 Gryllus bim
13	10	23.3	8	1	HIF1_PERAM	P04548 Periplaneta
14	10	23.3	8	1	HTP_TENMO	P25419 Tenebrio mo
15	10	23.3	8	1	LCK2_LEUMA	P21111 Leucophaea
16	10	23.3	8	1	LCK5_LEUMA	P13097 Leucophaea
17	10	23.3	8	1	LCK7_LEUMA	P13435 Leucophaea
18	10	23.3	8	1	RPCH_PANBO	P08959 Pandalus bo
19	10	23.3	8	1	UPA1_HUMAN	P30087 Homo sapien
20	10	23.3	9	1	FIBB_MACFU	P19835 Macaca fase
21	10	23.3	9	1	ISOF_CYPCA	P31239 Cyprinus ca
22	10	23.3	9	1	NEP_HV1Z8	P12481 Human immun
23	10	23.3	9	1	OXYT_RAJCL	P43274 Raja clavat
24	10	23.3	9	1	ULAE_HUMAN	P43251 Homo sapien
25	9	20.9	4	1	FEKA_ANTEL	P58756 Anthopleura
26	9	20.9	7	1	EI05_LITRU	P82250 Litoria rub
27	9	20.9	7	1	LANC_CAKUT	P30990 Natronobact
28	9	20.9	8	1	LCK1_LEUMA	P21340 Leucophaea
29	9	20.9	9	1	HDFU_KLEAE	P12981 Klebsiella
30	9	20.9	9	1	IAL1_PICJA	P17440 Pichia jadi
31	8	18.6	5	1	UC22_MAIZE	P80628 Zea mays (m
32	8	18.6	6	1	OVM_LEPVE	P14785 Leptinotars
33	8	18.6	7	1	MNP1_LEPDE	P42094 Leptinotars
34	8	18.6	8	1	ALC_YARMA	P81809 Yarrowia ma
35	8	18.6	8	1	644K_PORCC	P81886 Porphyromon
36	8	18.6	8	1	FUSS_FUSSO	P81917 Fusarium so
37	8	18.6	8	1	LCK4_LEUMA	P21343 Leucophaea
38	8	18.6	8	1	LIMS_STAEP	P48811 Staphylococ
39	8	18.6	8	1	NS3_MYCTU	P41252 Mycobacteri
40	8	18.6	9	1	CONO_CONSI	P55487 Conus striu
41	8	18.6	9	1	DNE1_LOCMI	P16899 Locusta mig
42	8	18.6	9	1	DSIP_RABIT	P11158 Oryctolagus
43	8	18.6	9	1	FAK3_CALVO	P41808 Calliphora
44	8	18.6	9	1	FAR6_CALVO	P41861 Calliphora
45	8	18.6	9	1	NEUX_HUMAN	P04277 Homo sapien

ALIGNMENTS

RESULT 1
IGAO_DACDE
ID IGAO_DACDE STANDARD; PRT: 7 AA.
AC P06294;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DT 01-OCT-1994 (Rel. 30, Last annotation update)
DE Galactose oxidase inhibitor.
OS Dactylium dendroides (Cladobotryum dendroides).
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Hypocreales; Hypocreaceae; Hypomyces.
OX NCBI_TaxID=5132;
RN [1]
RP SEQUENCE.
RA Avigad G., Markus Z.;
RT "Identification of a peptide inhibitor of galactose oxidase from
RT Dactylium dendroides.";
RL Fed. Proc. 31:447-447(1972).
CC -!- FUNCTION: BINDS ONE COPPER ION PER MOLECULE BUT DOES NOT BIND THE
CC GALACTOSE OXIDASE APOENZYME. IT MAY INACTIVATE THE ENZYME BY
CC BINDING TO ITS PROSTHETIC COPPER GROUP.
DR PIR; A01841; XEYDGD.
KW Copper; Metalloenzyme inhibitor.
SQ SEQUENCE 7 AA; 706 MW; 75BB01A456D871D0 CRC64;

Query Match 30.2%; Score 13; DB 1; Length 7;
Best Local Similarity 66.7%; Pred. No. 1e+05;
Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 7 NTQ 9
|||
Db 4 NIE 6

RESULT 2
IAL3_PICJA
ID IAL3_PICJA STANDARD; PRT: 9 AA.
AC P17441;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Transaldolase III (EC 2.2.1.2) (Fragment).
OS Pichia jadinii (Yeast) (Candida utilis).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Pichia.
OX NCBI_TaxID=4903;
RN [1]
RP SEQUENCE.
RX MEDLINE=75145197; PubMed=1092208;
RA Tsolas O., Sun S.C.;
RT "Isolation of a peptide containing a histidinyl cysteinyl sequence
RT from the active center of transaldolase.";
RL Arch. Biochem. Biophys. 167:525-533(1975).
CC -!- FUNCTION: TRANSALDOLASE IS IMPORTANT FOR THE BALANCE OF
CC METABOLITES IN THE PENTOSE PHOSPHATE PATHWAY.

```
CC     !  CATALYTIC ACTIVITY: Sedoheptulose 7 phosphate + D-glyceraldehyde
CC        3 phosphate = D-erythrose 4 phosphate + D-fructose 6 phosphate.
CC     !  PATHWAY: Pentose phosphate pathway; non-oxidative part.
CC     !  SIMILARITY: BELONGS TO THE TRANSALDOLASE FAMILY, SUBFAMILY 1.
DR   PIR; A11497; A11497.
DR   InterPro; IPR001585; Transaldolase.
DR   ProSite; PS00958; TRANSALDOLASE_2; PARTIAL.
DR   ProSite; PS01054; TRANSALDOLASE_1; PARTIAL.
KW   Transferase; Pentose shunt.
FT   NON_TER       1       1
FT   NON_TER       9       9
SQ   SEQUENCE   9 AA;  1033 MW;  325A31A14EB1E058 CRC64;


  Query Match             27.9%;  Score 12;  DB 1;  Length 9;
  Best Local Similarity   60.0%;  Pred. No. 1e+05;
  Matches    3;  Conservative    0;  Mismatches    2;  Indels    1;  Gaps    0;

Qy       4 IASNI 8
           ||
Db       5 IRSNI 9


RESULT   4
UNC5_CAEEL
ID   UN5_CAEEL      STANDARD;      PRT;     9 AA.
AC   P01451;
DT   15-JUL-1988 (Rel. 06, Created)
DT   15-JUL-1988 (Rel. 06, Last sequence update)
DT   15-JUL-1988 (Rel. 06, Last annotation update)
DE   Unknown protein 5B.6 from 2L pore (Fragment).
OS   Clostridium pasteurianum.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
OC   Clostridium.
OX   NCBI_TaxID=1501;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=88026571; PubMed=3143872;
RA   Pleau J.-M., Dardenne M., Blouquit Y., Bach J.-F.;
RT   "Structural study of circulating thymic factor: a peptide isolated
RT   from pig serum. II. Amino acid sequence.";
RL   J. Biol. Chem. 262:8045-8047(1987).
CC   -!- MISCELLANEOUS: THE BIOLOGICAL SOURCE(S) AND PHYSIOLOGICAL
CC       ACTIVITIES OF THIS PEPTIDE HAVE NOT BEEN DETERMINED.
DR   PIR; A01523; YFPG.
FT   MOD_RES       1       1       PYRROLIDONE CARBOXYLIC ACID.
SQ   SEQUENCE   9 AA;  876 MW;  1500A9B7B6B6C5B5A4D CRC64;


  Query Match             25.6%;  Score 11;  DB 1;  Length 9;
  Best Local Similarity   50.0%;  Pred. No. 1e+05;
  Matches    3;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       2 SKIASN 7
           :  ||
Db       4 SQGGSN 9


RESULT   5
THYF_PIG
ID   THYF_PIG       STANDARD;      PRT;     9 AA.
AC   P01455;
DT   21-JUL-1986 (Rel. 01, Created)
DT   21-JUL-1986 (Rel. 01, Last sequence update)
DT   21-JUL-1986 (Rel. 01, Last annotation update)
DE   Thymic factor.
OS   Sus scrofa (Pig).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX   NCBI_TaxID=9823;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=78026571; PubMed=914862;
RA   Pleau J.-M., Dardenne M., Blouquit Y., Bach J.-F.;
RT   "Structural study of circulating thymic factor: a peptide isolated
RT   from pig serum. II. Amino acid sequence.";
RL   J. Biol. Chem. 252:8045-8047(1977).
CC   -!- MISCELLANEOUS: THE BIOLOGICAL SOURCE(S) AND PHYSIOLOGICAL
CC       ACTIVITIES OF THIS PEPTIDE HAVE NOT BEEN DETERMINED.
DR   PIR; A01523; YFPG.
FT   MOD_RES       1       1       PYRROLIDONE CARBOXYLIC ACID.
SQ   SEQUENCE   9 AA;  876 MW;  1500A9B7B6B6C5B5A4D CRC64;


  Query Match             25.6%;  Score 11;  DB 1;  Length 9;
  Best Local Similarity   50.0%;  Pred. No. 1e+05;
  Matches    3;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       2 SKIASN 7
           :  ||
Db       4 SQGGSN 9


RESULT   6
FYRI_ANTEL
ID   FYRI_ANTEL     STANDARD;      PRT;     4 AA.
AC   P58706;
DT   01-MAR-2002 (Rel. 41, Created)
DT   01-MAR-2002 (Rel. 41, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
DE   Antho RFamide I [Contains: Antho RFamide II].
OS   Anthopleura elegantissima (Sea anemone).
OC   Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia; Actiniaria;
OC   Nynantheae; Actiniidae; Anthopleura.
OX   NCBI_TaxID=6110;
RN   [1]
RP   SEQUENCE.
RX   PubMed=18210965;
RA   Nothacker H.-P., Kinehart K.L. Jr., McFarlane I.R.,
RA   Grimmelikhuijzen C.J.P.;
RT   "Isolation of two novel neuropeptides from sea anemones: the unusual,
```

```
RT   biologically active L-3 phenyllactyl-Tyr-Arg-Ile-NH2 and its
RT   desphenyllactyl fragment Tyr-Arg-Ile-NH2.";
RL   Peptides 12:1169-1173(1991).
RN   [2]
RP   FUNCTION.
RX   PubMed 8397415;
RA   McFarlane I.D., Bradman D., Nothacker H.-P., Grimmelikhuijzen C.J.I.,
RT   "The expansion behaviour of sea anemones may be coordinated by two
RT   inhibitory neuropeptides, Antho-RFamide and Antho-kinamide.";
RL   Proc. R. Soc. Lond., B, Biol. Sci. 253:183-188(1993).
CC   -!- FUNCTION: Inhibits spontaneous contractions in several muscle
CC       strips. May be involved in the expansion phase of tactile
CC       behaviour in sea anemones.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- TISSUE SPECIFICITY: Neuron-specific.
KW   Neuropeptide; Amidation.
FT   CHAIN         1       4       ANTHO-RFAMIDE I.
FT   CHAIN         2       4       ANTHO-RFAMIDE II.
FT   MOD_RES       1       1       L-3-PHENYLLACTYL.
FT   MOD_RES       4       4       AMIDATION.
SQ   SEQUENCE   4 AA;  598 MW;  6044B859A0000000 CRC64;


  Query Match             23.8%;  Score 10;  DB 1;  Length 4;
  Best Local Similarity   50.0%;  Pred. No. 1e+05;
  Matches    2;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 FSKI 4
           : ||
Db       1 FYRI 4


RESULT   7
PSK_DAUCA
ID   PSK_DAUCA      STANDARD;      PRT;     5 AA.
AC   P58261;
DT   15-OCT-2001 (Rel. 40, Created)
DT   15-OCT-2001 (Rel. 40, Last sequence update)
DT   15-OCT-2001 (Rel. 40, Last annotation update)
DE   Phytosulfokine alpha (PSK-alpha) [Contains: Phytosulfokine-beta (PSK
DE   beta)].
OS   Daucus carota (Carrot).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC   Asteridae; euasterids II; Apiales; Apiaceae; Daucus.
OX   NCBI_TaxID=4039;
RN   [1]
RP   SEQUENCE, AND IDENTIFICATION BY MASS SPECTROMETRY.
RC   STRAIN=cv. US Harumakigosun;
RX   MEDLINE 20212743; PubMed 10750705;
RA   Hanai H., Matsuno T., Yamamoto M., Matsubayashi Y., Kobayashi I.,
RA   Kamada H., Sakagami Y.;
RT   "A secreted peptide growth factor, phytosulfokine, active in a
RT   stimulatory factor of carrot somatic embryo formation.";
RL   Plant Cell Physiol. 41:27-32(2000).
CC   -!- FUNCTION: IN PRESENCE OF 2,4 D, STIMULATES PROLIFERATION OF THE
CC       CELLS, BUT DOES NOT STIMULATE DIFFERENTIATION INTO THE 3 MAIN
CC       EMBRYOS.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- PTM: SULFATION IS IMPORTANT FOR ACTIVITY AND FOR THE BINDING TO A
CC       PUTATIVE MEMBRANE RECEPTOR (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE PHYTOSULFOKINE FAMILY.
KW   Growth factor; Sulfation.
FT   PEPTIDE       1       4       PHYTOSULFOKINE BETA.
FT   MOD_RES       1       1       SULFATION.
FT   MOD_RES       3       3       SULFATION.
SQ   SEQUENCE   5 AA;  497 MW;  7A73BE0B1B63A3A5 CRC64;


  Query Match             23.8%;  Score 10;  DB 1;  Length 5;
  Best Local Similarity   100.0%;  Pred. No. 1e+05;
  Matches    2;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       8 TQ 9
           ||
Db       4 TQ 5


RESULT   8
IPIS_CANFA
ID   IPIS_CANFA     STANDARD;      PRT;     7 AA.
AC   P54714;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   15-JUL-1998 (Rel. 36, Last annotation update)
DE   Triosephosphate isomerase (EC 5.3.1.1) (TIM) (Fragment)
GN   IPI1.
OS   Canis familiaris (dog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX   NCBI_TaxID=9615;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Heart;
RX   MEDLINE 98163340; PubMed 9564812;
RA   Dunn M.J., Corbett J.M., Wheeler C.H.;
RT   "HSC-2DPAGE and the two-dimensional gel electrophoresis database of
RT   dog heart proteins.";
RL   Electrophoresis 18:2795-2802(1997).
CC   -!- CATALYTIC ACTIVITY: D-glyceraldehyde 3-phosphate = glycerone
CC       phosphate.
CC   -!- PATHWAY: PLAYS AN IMPORTANT ROLE IN SEVERAL METABOLIC PATHWAYS.
CC   -!- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE TRIOSEPHOSPHATE ISOMERASE FAMILY.
DR   HSC-2DPAGE; P54714; DOG.
DR   InterPro; IPR000652; Trioseph_isomerase.
DR   PROSITE; PS00171; TIM; PARTIAL.
KW   Isomerase; Glycolysis; Gluconeogenesis; Fatty acid biosynthesis;
KW   Pentose shunt.
FT   NON_TER       1       1
FT   NON_TER       5       5
SQ   SEQUENCE   5 AA;  550 MW;  6444486209A0C000 CRC64;


  Query Match             23.8%;  Score 10;  DB 1;  Length 5;
  Best Local Similarity   25.0%;  Pred. No. 1e+05;
  Matches    1;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       4 IASN 7
           : :|
Db       2 VINN 5


RESULT   9
LOKI_LOCMI
ID   LOKI_LOCMI     STANDARD;      PRT;     9 AA.
AC   P41491;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   15-JUL-1998 (Rel. 36, Last annotation update)
DE   Locustakinin I.
OS   Locusta migratoria (Migratory locust).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Orthopteroidea; Orthoptera; Caelifera;
OC   Acridomorpha; Acridoidea; Acrididae; Locusta.
OX   NCBI_TaxID=7004;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Corpora cardiaca;
RX   MEDLINE 92262851; PubMed 1585017;
RA   Schoofs L., Holman G.M., Proost P., Van Damme J., Hayes T.K.,
RA   de Loof A.;
RT   "Locustakinin, a novel myotropic peptide from Locusta migratoria:
RT   isolation, primary structure and synthesis.";
RL   Regul. Pept. 37:49-57(1992).
CC   -!- FUNCTION: MYOTROPIC PEPTIDE. MAY BE INVOLVED IN THE STIMULATION
```

```
RA   Touze F., Riccardi F., Morin C., Dumerval C., Buet J. C.,
RA   Pernollet J.-C., Zivy M., de Vienne D.;
RT   "The maize two dimensional gel protein database: towards an integrated
RT   genome analysis program.";
RL   Theor. Appl. Genet. 93:997-1005(1996).
CC   -!- MISCELLANEOUS: ON THE 2D GEL THE DETERMINED PI OF THIS UNKNOWN
CC       PROTEIN IS: 6.0, ITS MW IS: 30.0 kDa.
DR   Maize-2DPAGE; P80640; COLEOPTILE.
DR   MaizeDB; 123956; -.
FT   NON_TER       1      1
FT   NON_TER       7      7
SQ   SEQUENCE   7 AA;  665 MW;  61C1B5B33DC1B5D0 CRC64;


  Query Match             24.3%;  Score 10;  DB 1;  Length 7;
  Best Local Similarity   50.0%;  Pred. No. 1e+05;
  Matches    2;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       5 ASNT 8
           | :|
Db       3 AKST 6


RESULT  12
AKHG_GRYBI
ID   AKHG_GRYBI     STANDARD;      PRT;     8 AA.
AC   P14086;
DT   01-JAN-1990 (Rel. 18, Created)
DT   01-FEB-1994 (Rel. 28, Last sequence update)
DT   01-FEB-1994 (Rel. 28, Last annotation update)
DE   Adipokinetic hormone G (AKH-G) (Ro II).
OS   Gryllus bimaculatus (Two-spotted cricket), and
OS   Romalea microptera (Lubber grasshopper).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Orthopteroidea; Orthoptera; Ensifera; Gryllidae;
OC   Gryllinae; Gryllus.
  X   NCBI_TaxID=6999, 7007;
RN   [1]
RP   SEQUENCE.
RC   SPECIES=G.bimaculatus; TISSUE=Corpora cardiaca;
RX   MEDLINE=88106553; PubMed=3226016;
RA   Gaede G., Rinehart K.L. Jr.;
RT   "Primary sequence analysis by fast atom bombardment mass spectrometry
RT   of a peptide with adipokinetic activity from the corpora cardiaca of
RT   the cricket Gryllus bimaculatus.";
RL   Biochem. Biophys. Res. Commun. 149:908-914(1987).
RN   [2]
RP   SEQUENCE.
RC   SPECIES=R.microptera; TISSUE=Corpora cardiaca;
RX   MEDLINE=89145002; PubMed=3226948;
RA   Gaede G., Hilbich C., Beyreuther K., Rinehart K.L. Jr.;
RT   "Sequence analyses of two neuropeptides of the AKH/RCP family from
RT   the lubber grasshopper, Romalea microptera.";
RL   Peptides 9:681-688(1988).
CC   -!- FUNCTION: THIS HORMONE, RELEASED FROM CELLS IN THE CORPORA
CC       CARDIACA AFTER THE BEGINNING OF FLIGHT, CAUSES RELEASE OF
CC       DIGLYCERIDES FROM THE FAT BODY AND THEN STIMULATES THE FLIGHT
CC       MUSCLES TO USE THESE DIGLYCERIDES AS AN ENERGY SOURCE.
CC   -!- SIMILARITY: BELONGS TO THE AKH/RCP HORMONE FAMILY.
DR   PIR; A28004; A28004.
DR   InterPro; IPR002047; AKH.
DR   PROSITE; PS00256; AKH; 1.
KW   Neuropeptide; Amidation; Flight.
FT   MOD_RES       1      1       PYRROLIDONE CARBOXYLIC ACID.
FT   MOD_RES       8      8       AMIDATION.
SQ   SEQUENCE   8 AA;  938 MW;  86C786B5B09C452D6 CRC64;


  Query Match             24.3%;  Score 10;  DB 1;  Length 8;
  Best Local Similarity   100.0%;  Pred. No. 1e+05;
  Matches    2;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 FS 2
```

```
Db       4 FS 5


RESULT  13
HFF_PERAM
ID   HFF_PERAM      STANDARD;      PRT;      8 AA.
AC   P25454;
DT   01-NOV-1992 (Rel. 25, Created)
DT   01-FEB-1994 (Rel. 28, Last sequence update)
DT   01-FEB-1994 (Rel. 28, Last annotation update)
DE   Hypertrehalosaemic factor I (Neuropeptide M-I) (Periplaneta M-I)
DE   (Pea-CAH I) (Leu-CC-I) (Hypertrehalosaemic neuropeptide I).
OS   Periplaneta americana (American cockroach),
OS   Leptinotarsa decemlineata (Colorado potato beetle), and
OS   Blatta orientalis (oriental cockroach).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Orthopteroidea; Dictyoptera; Blattaria;
OC   Blattidea; Blattidae; Periplaneta.
OX   NCBI_TaxID=6978; 7539, 6976;
RN   [1]
RP   SEQUENCE.
RC   SPECIES=P.americana;
RX   MEDLINE=85046550; PubMed=6548628;
RA   Witten J.L., Schaffer M.H., O'Shea M., Cook J.C., Hemlinat M.E.,
RA   Rinehart K.L. Jr.;
RT   "Structures of two cockroach neuropeptides assigned by fast atom
RT   bombardment mass spectrometry.";
RL   Biochem. Biophys. Res. Commun. 124:350-358(1984).
RN   [2]
RP   SEQUENCE.
RC   SPECIES=P.americana;
RX   MEDLINE=84299279; PubMed=6591205;
RA   Scarborough R.M., Rivier J., Kaiser E.T., Krantz D.E., Miller C.A.,
RA   Miller C.A., Schooley D.A.;
RT   "Isolation and primary structure of two peptides with
RT   cardiovasculatory and hyperglycemic activity from the corpora
RT   cardiaca of Periplaneta americana.";
RL   Proc. Natl. Acad. Sci. U.S.A. 81:5575-5579(1984).
RN   [3]
RP   SEQUENCE.
RC   SPECIES=L.decemlineata; TISSUE=Corpora cardiaca;
RX   MEDLINE=90250053; PubMed=2576128;
RA   Gaede G., Kellner R.;
RT   "The metabolic neuropeptides of the corpus cardiacum from the potato
RT   beetle and the American cockroach are identical.";
RL   Peptides 10:1287-1289(1989).
RN   [4]
RP   SEQUENCE.
RC   SPECIES=B.orientalis; TISSUE=Corpora cardiaca;
RX   MEDLINE=90250059; PubMed=2340112;
RA   Gaede G., Rinehart K.L. Jr.;
RT   "Primary structures of hypertrehalosaemic neuropeptides isolated from
RT   the corpora cardiaca of the cockroaches Leucophaea maderae,
RT   Gromphadorhina portentosa, Blattella germanica and Blatta orientalis
RT   and of the stick insect Extatosoma tiaratum assigned by fast
RT   atom bombardment mass spectrometry.";
RL   Biol. Chem. Hoppe-Seyler 371:345-354(1990).
CC   -!- FUNCTION: HYPERTREHALOSAEMIC FACTORS ARE NEUROSECRETORIES THAT
CC       ELEVATE THE LEVEL OF TREHALOSE IN THE HEMOLYMPH (TREHALOSE IS
CC       THE MAJOR CARBOHYDRATE IN THE HEMOLYMPH OF INSECTS).
CC   -!- SIMILARITY: BELONGS TO THE AKH / HRTH / RPCH FAMILY.
DR   PIR; A61259; A61259.
DR   PIR; S08095; S08095.
DR   PIR; A43982; A39927.
DR   PIR; A44966; A44966.
DR   InterPro; IPR002047; AKH.
DR   PROSITE; PS00256; AKH; 1.
KW   Neuropeptide; Amidation.
FT   MOD_RES       1       1       PYRROLIDONE CARBOXYLIC ACID.
FT   MOD_RES       8       8       AMIDATION.
SQ   SEQUENCE   8 AA;  991 MW;  86745775B9C452D6 CRC64;

  Query Match             23.3%;  Score 10;  DB 1;  Length 8;
  Best Local Similarity   100.0%;  Prod. No. 1e+05;
  Matches    2;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 FS 2
           ||
Db       4 FS 5


RESULT  14
HFF_TENMO
ID   HFF_TENMO      STANDARD;      PRT;      8 AA.
AC   P25419;
DT   01-MAY-1992 (Rel. 22, Created)
DT   01-FEB-1994 (Rel. 28, Last sequence update)
DT   01-FEB-1994 (Rel. 28, Last annotation update)
DE   Hypertrehalosaemic factor (HOTH) (Hypertrehalosaemic neuropeptide).
OS   Tenebrio molitor (Yellow mealworm), and
OS   Zophobas rugipes.
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Coleoptera; Polyphaga;
OC   Cucujiformia; Tenebrionidae; Tenebrio.
OX   NCBI_TaxID=7067, 7075;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Corpora cardiaca;
RX   MEDLINE=90341081; PubMed=2381871;
RA   Gaede G., Rosinski G.;
RT   "The primary structure of the hypertrehalosemic neuropeptide from
RT   tenebrionid beetles, a novel member of the AKH/RPCH family.";
RL   Peptides 11:455-459(1990).
CC   -!- FUNCTION: HYPERTREHALOSAEMIC FACTORS ARE NEUROSECRETORIES THAT
CC       ELEVATE THE LEVEL OF TREHALOSE IN THE HEMOLYMPH (TREHALOSE IS
CC       THE MAJOR CARBOHYDRATE IN THE HEMOLYMPH OF INSECTS).
CC   -!- SIMILARITY: BELONGS TO THE AKH / HRTH / RPCH FAMILY.
DR   PIR; A43976; A43976.
DR   PIR; B43976; B43976.
DR   InterPro; IPR002047; AKH.
DR   PROSITE; PS00256; AKH; 1.
KW   Neuropeptide; Amidation.
FT   MOD_RES       1       1       PYRROLIDONE CARBOXYLIC ACID.
FT   MOD_RES       8       8       AMIDATION.
SQ   SEQUENCE   8 AA;  1005 MW;  86745775B9C4A7A6 CRC64;

  Query Match             23.3%;  Score 10;  DB 1;  Length 8;
  Best Local Similarity   100.0%;  Prod. No. 1e+05;
  Matches    2;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 FS 2
           ||
Db       4 FS 5


RESULT  15
LCK2_LEUMA
ID   LCK2_LEUMA      STANDARD;      PRT;      8 AA.
AC   P21141;
DT   01-MAY-1991 (Rel. 18, Created)
DT   01-MAY-1991 (Rel. 18, Last sequence update)
DT   01-MAY-1991 (Rel. 18, Last annotation update)
DE   Leucokinin II (L II).
OS   Leucophaea maderae (Madeira cockroach).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Orthopteroidea; Dictyoptera; Blattaria;
OC   Blaberoidea; Blaberidae; Leucophaea.
OX   NCBI_TaxID=6988;
RN   [1]
RP   SEQUENCE, AND SYNTHESIS.
RC   TISSUE=Head;
RA   Holman G.M., Cook B.J., Nachman R.J.;
RT   "Isolation, primary structure and synthesis of two new peptides
```

RT [illegible] peptides moderate members of a new family of
RT [illegible] myotropins.";
RL [illegible] Physiol. 84:205-211(1986).
CC -!- FUNCTION: THIS CEPHALOMYOTROPIC PEPTIDE STIMULATES CONTRACTILE
CC ACTIVITY OF CockROACH PROVENTRIUM (HINDGUT).
CC -!- SIMILARITY TO THE OTHER LEUCOKININS.
KW Neuropeptide; Amidation.
FT MOD_RES [illegible] AMIDATION.
SQ [illegible] AA; [illegible] MW; [illegible] CRC64;

Query Match 23.9%; Score 10; DB 1; Length 7;
Best Local Similarity 100.0%; Pred. No. [illegible];
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PS 2
:
Db 1 [illegible]

Search completed: August 26, 2002, 06:40:44
Job time [illegible]

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

SMPstein: protein search, using sw model

Run on: August 26, 2002, 06:39:47 ; Search time 15.06 Seconds
(without alignments)
11.455 Million cell updates/sec

Title: US-09-744-282-4
Perfect score: 43
Sequence: 1 FSKIASNTQ 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0
Maximum DB seq length: 9

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_19:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	DB	ID	Description
1	18	41.9	9	12 Q88953	Q88953 Vaccinia vi
2	17	39.5	8	2 O87471	O87471 haemophilus
3	17	39.5	8	11 Q62527	Q62527 mus spretus
4	15	34.9	8	12 Q91U21	Q91U21 Influenza a
5	15	34.9	8	12 Q91U19	Q91U19 Influenza a
6	14	32.6	7	2 P73884	P73884 Cyanobacter
7	14	32.6	9	6 Q9IRS0	Q9IRS0 Oryctolagus
8	13	30.2	8	2 Q47273	Q47273 Escherichia
9	13	30.2	8	4 Q16428	Q16428 Homo sapien
10	13	30.2	8	4 O60774	O60774 Homo sapien
11	13	30.2	8	6 Q9MY15	Q9MY15 Potato pyema
12	13	30.2	9	15 Q85710	Q85710 Ovis Sarcom
13	12	27.9	8	2 Q92E29	Q92E29 Ehrlichia ph
14	12	27.9	8	2 Q44464	Q44464 Thiobacca
15	12	27.9	8	3 P87225	P87225 Saccharomyc
16	12	27.9	8	6 O18854	O18854 Canis famili
17	12	27.9	8	10 Q9SAY7	Q9SAY7 Diospyros k
18	12	27.9	8	12 Q84176	Q84176 Orf virus (
19	12	27.9	8	13 P82282	P82282 Limnodynast
20	12	27.9	9	4 Q9B522	Q9B522 Homo Sapien
21	12	27.9	9	5 Q91VF1	Q91VF1 Trypanosoma
22	11	25.6	7	2 O54228	O54228 Sphingomona
23	11	25.6	8	2 Q9F517	Q9F517 Clostridium
24	11	25.6	8	2 O68485	O68485 Klebsiella
25	11	25.6	8	2 Q49534	Q49534 Mycoplasma
26	11	25.6	8	3 O13591	O13591 Saccharomyc
27	11	25.6	8	4 Q15869	Q15869 Homo Sapien
28	11	25.6	8	4 Q15895	Q15895 Homo Sapien
29	11	25.6	8	7 Q29801	Q29801 Homo Sapien
30	11	25.6	8	7 Q29797	Q29797 Homo Sapien
31	11	25.6	8	7 Q29812	Q29812 Homo Sapien
32	11	25.6	8	8 Q9I2W0	Q9I2W0 Saccharomyc
33	11	25.6	8	12 Q66807	Q66807 Human echov
34	11	25.6	8	13 P82083	P82083 Limnodynast
35	11	25.6	8	13 Q90ZV5	Q90ZV5 Tulipa fore
36	11	25.6	9	2 Q08810	Q08810 Borrelia af
37	11	25.6	9	2 Q46179	Q46179 Clostridium
38	11	25.6	9	2 Q98792	Q98792 Borrelia bu
39	11	25.6	9	2 O51463	O51463 Borrelia ta
40	11	25.6	9	3 Q90K18	Q90K18 Seleteriam
41	11	25.6	9	4 Q60QA5	Q60QA5 Homo Sapien
42	11	25.6	9	5 Q9VV82	Q9VV82 Drosophila
43	11	25.6	9	6 P82926	P82926 Bos Taurus
44	11	25.6	9	8 Q95GN1	Q95GN1 Petaloto am
45	11	25.6	9	15 Q20C43	Q20C43 Human immun

ALIGNMENTS

RESULT 1
Q88953
ID Q88953 PRELIMINARY; PRT; 9 AA.
AC Q88953;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE SERPINS (FRAGMENT).
GN B13R/SPI-2.
OS Vaccinia virus.
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC Orthopoxvirus.
OX NCBI_TaxID=10245;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95133144; PubMed=7831769;
RA Kettle S., Blake N.W., Law K.M., Smith G.L.;
RT "Vaccinia virus serpins B13R (SPI-2) and B22R (SPI-1) encode M(r) 38,
RT and 40K, intracellular polypeptides that do not affect virus virulence
RT in a murine intranasal model.";
RL Virology 206:136-147(1995).
DR EMBL; S75133; AAC60746.1; -.
FT NON_TER 9 9
SQ SEQUENCE 9 AA; 1081 MW; 2E6A1AC5ED4E4C7A CRC64;

Query Match 41.9%; Score 18; DB 12; Length 9;
Best Local Similarity 66.7%; Pred. No. 5.2e-05;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FSKIAS 6
F+ :|||
Db 4 FREIAS 9

RESULT 2
O87471
ID O87471 PRELIMINARY; PRT; 8 AA.
AC O87471;

```
DT   01-NOV-1998 (TrEMBLrel. 08, Created)
DT   01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT   01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE   HlpA (FRAGMENT).
GN   HlpA.
OS   Haemophilus influenzae.
OC   Bacteria; Proteobacteria; Gamma subdivision; Pasteurellaceae;
OC   Haemophilus.
OX   NCBI_TaxID=727;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN PAGAN (EIA);
RX   MEDLINE=98099808; PubMed=9721513;
RA   Mhlanga-Mutangadura T., Morlin G., Smith A.L., Eisenstark A.;
RA   Gotschlich M.;
RT   "Evidence that the major pilus gene cluster of Haemophilus
RT   influenzae.";
RL   J. Bacteriol. 180:4693-4703(1998).
DR   EMBL; AF017762; AAC45830.1; -.
FT   NON_TER       1       1
SQ   SEQUENCE   8 AA;  876 MW;  DAB44451A727A329 CRC64;


  Query Match             89.5%;  Score 17;  DB 2;  Length 8;
  Best Local Similarity   50.0%;  Pred. No. 5.2e+05;
  Matches    4;  Conservative    2;  Mismatches    1;  Indels    1;  Gaps    0;

Qy       2 SKIASN 7
           |: :|
Db       1 SKEIGN 6


RESULT   4
Q9U5J7
ID   Q9U5J7       PRELIMINARY;      PRT;     8 AA.
AC   Q9U5J7;
DT   01-JUN-1998 (TrEMBLrel. 06, Created)
DT   01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT   01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
DE   TRANSTHYRETIN (PREALBUMIN) (FRAGMENT).
GN   TTR
OS   Mus spretus (Western wild mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10096;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN SPRET/EI;
RX   MEDLINE=94139082; PubMed=8044949;
RA   Ko M.S., Wang X., Horton J.H., Hagen M.D., Takahashi N., Maeyaki Y.,
RA   Nadeau J.H.;
RT   "Genetic mapping of 40 cDNA clones on the mouse genome by FCR.";
RL   Mamm. Genome 5:449-455(1994).
CC   -!- FUNCTION: CHYPOID HORMONE BINDING PROTEIN, RESPONSIBLE FOR TRANSPORTS
CC       CHYROXINE FROM THE BLOODSTREAM TO THE BRAIN.
CC   -!- SUBUNIT: HOMOTETRAMER.
DR   EMBL; U59096; AAB60461.1; -.
DR   MGD; MGI:96857; Ttr.
KW   Acute phase; Transport; Retinol binding; Vitamin A; Thyroid hormone.
FT   NON_TER       1       1
SQ   SEQUENCE   8 AA;  828 MW;  91D6C7B455A2D2C9 CRC64;


  Query Match             89.5%;  Score 17;  DB 11;  Length 8;
  Best Local Similarity   50.0%;  Prod. No. 5.2e+05;
  Matches    4;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 SKIASN 6
           |: :|
Db       1 SKEIGN 6


RESULT   5
```

```
Q91U21
ID   Q91U21       PRELIMINARY;      PRT;     8 AA.
AC   Q91U21;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   NUCLEAR EXPORT PROTEIN NS2 (FRAGMENT).
GN   NS.
OS   Influenza A virus (A/Hong Kong/156/97(H5N1)).
OC   Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
OC   Influenza virus A and B group; Influenza A viruses; Influenza A virus.
OX   NCBI_TaxID=155224;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN A/HONG KONG/156/97;
RA   Shaw M.W., Cooper L.A., Xu X., Thompson W.W., Krauss S.L., Guan Y.,
RA   Zhou N.N., Klimov A., Cox N.J., Webster R.G., Lim W., Shortridge K.F.,
RA   Subbarao K.;
RT   "Avian influenza A H5N1 and H9N2 viruses bearing a specific
RT   constellation of nonglycoprotein genes caused illness in humans.";
RL   Submitted (APR-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF256191; AAK49324.1; -.
FT   NON_TER       1       1
FT   NON_TER       8       8
SQ   SEQUENCE   8 AA;  869 MW;  D9DDD5A5A2D1A455 CRC64;


  Query Match             44.9%;  Score 15;  DB 12;  Length 8;
  Best Local Similarity  100.0%;  Prod. No. 5.6e+05;
  Matches    3;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       6 SNI 8
           |||
Db       1 SNI 3


RESULT   5
Q91U19
ID   Q91U19       PRELIMINARY;      PRT;     8 AA.
AC   Q91U19;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   NUCLEAR EXPORT PROTEIN NS2 (FRAGMENT).
GN   NS.
OS   Influenza A virus (A/Hong Kong/156/97(H5N1)).
OC   Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
OC   Influenza virus A and B group; Influenza A viruses; Influenza A virus.
OX   NCBI_TaxID=155224;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN A/HONG KONG/156/97;
RA   Shaw M.W., Cooper L.A., Xu X., Thompson W.W., Krauss S.L., Guan Y.,
RA   Zhou N.N., Klimov A., Cox N.J., Webster R.G., Lim W., Shortridge K.F.,
RA   Subbarao K.;
RT   "Avian influenza A H5N1 and H9N2 viruses bearing a specific
RT   constellation of nonglycoprotein genes caused illness in humans.";
RL   Submitted (APR-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF256192; AAK49326.1; -.
FT   NON_TER       1       1
FT   NON_TER       8       8
SQ   SEQUENCE   8 AA;  869 MW;  D9DDD5A5A2D1A455 CRC64;


  Query Match             44.9%;  Score 15;  DB 12;  Length 8;
  Best Local Similarity  100.0%;  Prod. No. 5.6e+05;
  Matches    3;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       6 SNI 8
           |||
Db       1 SNI 3
```

RESULT 6
P70804
ID P70804 PRELIMINARY; PRT; 7 AA.
AC P70804;
DT 01-FEB-1997 (TrEMBLrel. 02, Created)
DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT 31-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE ALGI PROTEIN (FRAGMENT).
GN ALGI.
OS Azotobacter vinelandii.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Azotobacter.
OX NCBI_TaxID=354;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN E;
RX MEDLINE 94427318; PubMed 8830682;
RA Rehm B.H.A., Ertesvag H., Valla S.;
RT "A new Azotobacter vinelandii mannuronan C-5-epimerase gene (algE5) is
RT part of an algE gene cluster physically organized in a manner similar
RT to that in Pseudomonas aeruginosa.";
RL J. Bacteriol. 178:5884-5889(1996).
DR EMBL; X87073; CAA61230.1; -.
FT NON_TER 1 1
SQ SEQUENCE 7 AA; 741 MW; 7DF4AFAFA2F7AEE CRC64;

Query Match 32.6%; Score 14; DB 2; Length 7;
Best Local Similarity 40.0%; Pred. No. 5.6e+05;
Matches 2; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 4 IASNI 8
:::|:
Db 3 VSSST 7

RESULT 7
Q9TRS0
ID Q9TRS0 PRELIMINARY; PRT; 9 AA.
AC Q9TRS0;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE CALCYCLIN ASSOCIATED PROTEIN, CAP50 CA2+/PHOSPHOLIPID BINDING PROTEIN
DE 1.7 FRAGMENT.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE.
RX MEDLINE 92250478; PubMed 1533622;
RA Tokumitsu H., Minami A., Minami R., Kobayashi R., Hidaka H.;
RT "A calcyclin associated protein is a newly identified member of the
RT Ca2+/phospholipid-binding proteins, annexin family.";
RL J. Biol. Chem. 267:8919-8924(1992).
SQ SEQUENCE 9 AA; 1010 MW; 64E419C44865B728 CRC64;

Query Match 32.6%; Score 14; DB 6; Length 9;
Best Local Similarity 33.3%; Pred. No. 5.6e+05;
Matches 2; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Qy 2 SKIASN 7
:|:::
Db 1 SESSIN 6

RESULT 8
Q47273
ID Q47273 PRELIMINARY; PRT; 8 AA.
AC Q47273;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE ORF13, ORF151, ORF56, ORF96, K03, ORF45, ..., 3' ... 3'. ANI NMPC
DE GENES.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN K-12;
RX MEDLINE 96196428; PubMed 8648624;
RA Maladi A.A., Sharples G.J., Mandal T.N., Lloyd R.G.;
RT "Holliday junction Resolvases encoded by homologous ruvA genes in
RT Escherichia coli K-12 and Phage 82.";
RL J. Mol. Biol. 257:561-573(1996).
DR EMBL; X92587; CAA63324.1; -.
SQ SEQUENCE 8 AA; 935 MW; FD4041A05B05B3C6 CRC64;

Query Match 30.2%; Score 13; DB 2; Length 8;
Best Local Similarity 50.0%; Pred. No. 5.6e+05;
Matches 3; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 2 PEIEN 7
:|:|:
Db 3 SEITN 8

RESULT 9
Q16428
ID Q16428 PRELIMINARY; PRT; 8 AA.
AC Q16428;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE DYSTROPHIN PROTEIN (FRAGMENT).
GN DYSTROPHIN.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE 96163501; PubMed 8566960;
RA Holder E., Maeda M., Bies R.D.;
RT "Expression and regulation of the dystrophin Purk-1 promoter in
RT human skeletal muscle, heart, and brain.";
RL Hum. Genet. 97:232-239(1996).
RN [2]
RP SEQUENCE FROM N.A.
RA Kitano T., Kobayakawa R., Saitou N.;
RT "Silver Project.";
RL Submitted (JAN-2000) to the EMBL/GenBank/DDBJ databases.
DR EMBL; S81419; AAB14063.1; -.
DR EMBL; AB037493; BAA90413.1; -.
FT NON_TER 8 8
SQ SEQUENCE 8 AA; 825 MW; F21865A5A2B65B6 CRC64;

Query Match 30.2%; Score 13; DB 4; Length 8;
Best Local Similarity 40.0%; Pred. No. 5.6e+05;
Matches 2; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 2 SKIAS 6
:::|:
Db 2 SEVSS 6

RESULT 10
O60773
ID O60773 PRELIMINARY; PRT; 8 AA.
AC O60773;

```
ID   ... ; ... ; ... (TrEMBLrel. 02, Created)
ID   ... ; ... ; ... (TrEMBLrel. 13, Last sequence update)
ID   ... ; ... ; ... (TrEMBLrel. 17, Last annotation update)
DE   ... (PARVALBUMIN) (FRAGMENT).
GN   ...VALB.
OS   ... Sapiens (Human).
OC   ... Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   ...
RP   SEQUENCE FROM N.A.
RA   ...;
RT   "...";
RL   Submitted (DEC 1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; ...; CAB50099.1; -.
FT   NON_TER       ...       1
SQ   SEQUENCE   8 AA;  853 MW;  9BBB1...  CRC64;


  Query Match             30.2%;  Score 13;  DB 1;  Length 8;
  Best Local Similarity   25.0%;  Pred. No. ...;
  Matches    2;  Conservative    3;  Mismatches    3;  Indels    ...  Gaps    ...;

Qy       ... EASIASN 8
             ...
Db       ... PSTIVAES 8


RESULT  ...
...
ID   ...       PRELIMINARY;      PRT;     8 AA.
AC   ...;
DT   ... (TrEMBLrel. 15, Created)
DT   ... (TrEMBLrel. 15, Last sequence update)
DT   ... (TrEMBLrel. 17, Last annotation update)
DE   ...
GN   ...
OS   ... troglodytes (chimpanzee), and
OS   ... gorilla (gorilla).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pongo.
OX   NCBI_TaxID=9600, 9598, 9593;
RN   ...
RP   SEQUENCE FROM N.A.
RC   ...
RA   Kajiwara H., Kobayakawa H., Sarkar N.;
RT   "...";
RL   Submitted (JAN 2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; ...; BAA90419.1; -.
DR   EMBL; ...; BAA90414.1; -.
DR   EMBL; ...; BAA90417.1; -.
SQ   SEQUENCE   8 AA;  825 MW;  F2B6...  CRC64;


  Query Match             30.2%;  Score 13;  DB 2;  Length 8;
  Best Local Similarity   40.0%;  Pred. No. ...;
  Matches    2;  Conservative    3;  Mismatches    2;  Indels    ...  Gaps    ...;

Qy       ... SKIAS 6
             ...
Db       ... SEVSS 6


RESULT  ...
...
ID   ...       PRELIMINARY;      PRT;     6 AA.
AC   ...;
DT   ... (TrEMBLrel. 01, Created)
DT   ... (TrEMBLrel. 01, Last sequence update)
DT   ... (TrEMBLrel. 08, Last annotation update)
DE   ... (FRAGMENT).
OS   ...
OC   Viruses; Retroviruses; Retroviridae; Alpharetrovirus.
OX   NCBI_TaxID=11866;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=84115080; PubMed=6319754;
RA   Lerner T.L., Hanafusa H.;
RT   "DNA sequence of the Bryan high titer strain of Rous sarcoma virus:
RT   Extent of env deletion and possible genealogical relationship with
RT   other viral strains.";
RL   J. Virol. 49:549-556(1984).
DR   EMBL; K03365; AAA42557.1; -.
FT   NON_TER       1       1
SQ   SEQUENCE   9 AA;  949 MW;  94AA144DDDD731AA CRC64;


  Query Match             30.2%;  Score 13;  DB 15;  Length 9;
  Best Local Similarity   50.0%;  Pred. No. ...;
  Matches    2;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       4 IASN 7
             :|:|
Db       4 LAAN 7


RESULT  13
Q9ZEZ9
ID   Q9ZEZ9      PRELIMINARY;      PRT;     8 AA.
AC   Q9ZEZ9;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-MAY-1999 (TrEMBLrel. 10, Last annotation update)
DE   ...
GN   LEUA.
OS   Buchnera aphidicola.
OG   Plasmid pBRel.
OC   Bacteria; Proteobacteria; gamma subdivision; Buchnera.
OX   NCBI_TaxID=9;
RN   [.]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=99028904; PubMed=9812361;
RA   Silva F.J., Van Ham R.C.H.J., Sabater B., Latorre A.;
RT   "Structure and evolution of the leucine plasmids carried by the
RT   endosymbiont (Buchnera aphidicola) from aphids of the family
RT   Aphididae.";
RL   FEMS Microbiol. Lett. 168:43-49(1998).
DR   EMBL; AJ006874; CAA07290.1; -.
KW   Lyase; Plasmid.
FT   NON_TER       8       8
SQ   SEQUENCE   8 AA;  917 MW;  EE2B03A2D67B446 CRC64;


  Query Match             27.9%;  Score 12;  DB 2;  Length 8;
  Best Local Similarity   66.7%;  Pred. No. ...;
  Matches    2;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       7 NIQ 9
             |:|
Db       2 NSQ 4


RESULT  14
Q44463
ID   Q44463      PRELIMINARY;      PRT;     8 AA.
AC   Q44463;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   AGROBACTERIUM TUMEFACIENS II PLASMID VIR14 AND VIR15 GENES
DE   (FRAGMENT).
OS   Rhizobiaceae group (rhizobacteria).
OG   Plasmid Ti.
OC   Bacteria; Proteobacteria; alpha subdivision.
OX   NCBI_TaxID=356;
RN   [1]
RP   SEQUENCE FROM N.A.
```

RX MEDLINE; 88[illegible]; PubMed 3058701;
RA Fuller S.J., Yanofsky M.F., Nester E.W.;
RT "Molecular characterization of the virB operon from Agrobacterium
RT tumefaciens.";
RL Nucleic Acids Res. 15:7503-7517(1987).
DR EMBL; X 6 45; CAA29439.1; -.
KW Plasmid.
FT NON_TER 8 8
SQ SEQUENCE 8 AA; 887 MW; F8F2C325B33861A6 CRC64;

Query Match 27.9%; Score 12; DB 2; Length 8;
Best Local Similarity 66.7%; Pred. No. 5.6e+05;
Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 SKI 4
:|:
Db 5 SKV 7

RESULT 15
P87225
ID P87225 PRELIMINARY; PRT; 8 AA.
AC P87225;
DT 01-JUL-1996 (TrEMBLrel. 04, Created)
DT 01-NOV-1996 (TrEMBLrel. 12, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE SINI (6 IEIN (FRAGMENT).
GN SINI OR YIL065W.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycetina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RA Wedler H., Wedler F., Schartl M., Wambutt R.;
RL Submitted (MAY-1996) to the EMBL/GenBank/DDBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA MIPS;
RL Submitted (MAY-1996) to the EMBL/GenBank/DDBJ databases.
DR EMBL; Z31[illegible]; CAA87518.2; -.
DR SGD; S0003988; GIN11.
FT NON_TER 1 1
SQ SEQUENCE 8 AA; 1319 MW; 4E21A9C449D587B6 CRC64;

Query Match 27.9%; Score 12; DB 8; Length 8;
Best Local Similarity 66.7%; Pred. No. 5.6e+05;
Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FSK 3
:|:
Db 7 FIK 9

Search completed: August 26, 2002, 06:39:48
Job time: 47 sec

GenCore version 4.5
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 26, 2002, 06:27:06 ; Search time 12.16 Seconds
(without alignments)
7.29 Million cell updates/sec

Title: US-09-744-282-4
Perfect score: 48
Sequence: 1 ESKIASNIQY 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0
Maximum DB seq length: 9

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: A_Geneseq_042902:*
1: /cgn2_6/ptodata/1/pubpep/A_Geneseq_042902.pep:*
2: /cgn2_6/ptodata/1/pubpep/B_Geneseq_042902.pep:*
3: /cgn2_6/ptodata/1/pubpep/C_Geneseq_042902.pep:*
4: /cgn2_6/ptodata/1/pubpep/D_Geneseq_042902.pep:*
5: /cgn2_6/ptodata/1/pubpep/E_Geneseq_042902.pep:*
6: /cgn2_6/ptodata/1/pubpep/F_Geneseq_042902.pep:*
7: /cgn2_6/ptodata/1/pubpep/backup_Geneseq_042902.pep:*
8: /cgn2_6/ptodata/1/pubpep/backup2_Geneseq_042902.pep:*
9: /cgn2_6/ptodata/1/pubpep/backup3_Geneseq_042902.pep:*
10: /cgn2_6/ptodata/1/pubpep/backup4_Geneseq_042902.pep:*
11: /cgn2_6/ptodata/1/pubpep/backup5_Geneseq_042902.pep:*
12: /cgn2_6/ptodata/1/pubpep/backup6_Geneseq_042902.pep:*
13: /cgn2_6/ptodata/1/pubpep/backup7_Geneseq_042902.pep:*
14: /cgn2_6/ptodata/1/pubpep/backup8_Geneseq_042902.pep:*
15: /cgn2_6/ptodata/1/pubpep/backup9_Geneseq_042902.pep:*
16: /cgn2_6/ptodata/1/pubpep/backup10_Geneseq_042902.pep:*
17: /cgn2_6/ptodata/1/pubpep/backup11_Geneseq_042902.pep:*
18: /cgn2_6/ptodata/1/pubpep/backup12_Geneseq_042902.pep:*
19: /cgn2_6/ptodata/1/pubpep/backup13_Geneseq_042902.pep:*
20: /cgn2_6/ptodata/1/pubpep/backup14_Geneseq_042902.pep:*
21: /cgn2_6/ptodata/1/pubpep/backup15_Geneseq_042902.pep:*
22: /cgn2_6/ptodata/1/pubpep/backup16_Geneseq_042902.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	48	100.0	9	17	AAR95074	Human cartilage gl
2	48	100.0	9	21	AAY69229	Peptide fragment o
3	27	56.2	6	21	AAY63423	Peptide motif from
4	21	43.8	7	21	AAY62448	Generic T4 epitope
5	21	43.8	8	13	AAR26082	Immunising peptide
6	21	43.8	8	14	AAR42960	Beta chain V8.1 seq
7	21	43.8	8	22	AAB08365	Immunogenic peptid
8	21	43.8	9	20	AAY53411	Bee Api epitope (a
9	21	43.8	9	20	AAY43133	Amino acid sequenc
10	21	43.8	9	20	AAY53449	bee Api derived pe
11	21	43.8	9	20	AAW97980	Human synaptonemal
12	21	43.8	9	21	AAY53238	SCL - BLA binding
13	21	43.8	9	22	AAB82788	Epitope of mycobac
14	21	43.8	9	22	AAM24552	Human MHC class I
15	20	41.5	7	21	AAY61286	Cadherin 6 peptid
16	20	41.5	7	21	AAY61427	Cadherin 7 peptid
17	20	41.5	7	21	AAY61429	Cadherin 7 peptid
18	20	41.5	7	21	AAY62663	T cadherin peptid
19	20	41.5	8	22	AAB98115	Human BNP associat
20	20	41.5	9	20	AAY45630	Immunogenic peptid
21	20	41.5	9	20	AAY46171	Immunogenic peptid
22	20	41.5	9	20	AAY60815	HLA A24 antigen f
23	20	41.5	9	22	AAM25072	Human MHC class I
24	20	41.5	9	22	AAM25154	human MHC associat
25	19	39.2	7	17	AAB69422	Antibody V-I disp
26	19	39.2	7	18	AAW22423	CDR2 from light ch
27	19	39.2	7	20	AAY55149	Anti-TL44 antibody
28	19	39.2	7	21	AAY60144	Chimeric anti-TL4
29	19	39.2	7	21	AAY60895	Human haematopoiet
30	19	39.2	7	21	AAY62250	Cadherin 14 V-II s
31	19	39.2	7	21	AAY63815	Desmoglein cell ad
32	19	39.2	7	22	AAE13478	Humanised Ab light
33	19	39.2	7	22	AAE13572	Humanised Ab light
34	19	39.2	7	22	AAG63248	Complement-activat
35	19	39.2	7	22	AAR86047	H. pylori database
36	19	39.2	7	22	AAR86079	H. pylori database
37	19	39.2	8	13	AAR26111	Beta chain V8.1 seq
38	19	39.2	8	20	AAY08150	Clotting factor V
39	19	39.2	8	20	AAY08158	Clotting factor V
40	19	39.2	8	20	AAY08159	Clotting factor V
41	19	39.2	8	20	AAW76797	Rat 20k VDI protei
42	19	41.2	8	21	AAY62291	Cadherin 14 precur
43	19	41.2	8	21	AAY62459	Cadherin 14 V-II s
44	19	41.2	8	21	AAY63458	Peptide of cadherin
45	19	41.2	8	21	AAY63484	Immunogenic peptid

ALIGNMENTS

RESULT 1
AAR95074
ID AAR95074 standard; peptide; 9 AA.
XX
AC AAR95074;
XX
DT [illegible] (first entry).
XX
DE Human cartilage glycoprotein 39 antigenic derived peptide (IV).
XX
KW Human; cartilage; glycoprotein 39; antigen; [illegible]
KW induction; E coli; tolerance; mediated; [illegible]
KW rheumatoid arthritis; disease model; diagnosis.
XX
OS Homo sapiens.
XX
PN WO9613517-A1.
XX
PD 09-MAY-1996.
XX
PF 25-OCT-1995; 95WO-EP04211.
XX
PR 07-APR-1995; 95EP-0200886.
PR 27-OCT-1994; 94EP-0203128.
XX
PA (ARLO-) AKZO NOBEL NV.
XX
PI Boots AMH, Verheijden GFM;
XX
DR WPI, 1996-239447/24.
XX
PT Peptide(s) derived from human cartilage gp 39 antigen - used
PT to induce specific E coli tolerance, esp. for treatment of
PT rheumatoid arthritis, also useful in predicting the risk

```
...  9 SeqA
XX
...  ... 1-MAY-2000; 97pp; English.
XX
...  ...treated' peptide, derived from the human cartilage glycoprotein
...  39 (HC gp 39) autoantigen, can be used to induce specific T-cell
...  ...erance to HC gp 39 in a patient with T-cell mediated cartilage
...  destruction, espc. rheumatoid arthritis (RA). It may also be used to
...  ...duce RA in mice, etc. (human disease models), and to diagnose the
...  presence of activated T-cells autoreactive with HC gp 39. Due to
...  ...city, patients suitable for tolerising treatment
...  Intranasally the peptide is given in a dose of 0.2 to 10 mg/kg
...  ...ously, or 10 to 50 micrograms/kg to produce a disease model. The
...  peptide has a specific tolerising effect on autoreactive T-cells,
...  ...ving other components of the immune system intact, and
...  ...ected to have few side effects. In an assay to determine
...  ...e peptides' binding to HLA DR molecules, specifically DRdw4
...  ...ve was associated with an increased risk of RA development, the
...  peptide gave a Dot blot D*(50) (microMolar) value of 0.08 compared
...  ...  ...  for the control peptide HHA(307-319).
XX
SQ   Sequence    9 AA;


  Query Match             100.0%;  Score 44;  DB 12;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 4.4e-05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 ESKIASNTQ 9
           |||||||||
Db       1 ESKIASNTQ 9


RESULT   2
AAY93423
ID   AAY93423 standard; peptide; 6 AA.
XX
AC   AAY93423;
XX
DT   04-SEP-2000  (first entry)
XX
DE   Peptide motif from a starch granule bound protein of rice.
XX
KW   Starch granule bound protein; rice; starch phosphorylation;
KW   transgenic plant; starch; cooking quality.
XX
OS   Oryza sativa.
XX
PN   WO200028052-A2.
XX
PD   18-MAY-2000.
XX
PF   05-NOV-1999;  99WO-EP008503.
XX
PR   09-NOV-1998;  98US-0107883.
XX
PA   (PLAN ) PLANTEC BIOTECHNOLOGIE GMBH.
XX
PI   Frohberg C;
XX
DR   WPI; 2000-376563/32.
XX
PT   Polynucleotide encoding R1 protein from rice useful for production of
PT   transgenic plants synthesizing modified starch.
XX
PS   Claim 2; Page 86; 99pp; English.
XX
CC   AAY93419-40 represent peptide motifs from a starch granule bound
CC   protein of rice. The protein is present in the plastids of plant
CC   cells, and is probably involved in the phosphorylation of starch.
CC   The polynucleotide sequence is useful for obtaining transgenic
CC   plants, especially rice, which are capable of synthesizing modified
CC   starch in the endogenous form (in reduced levels). Unlike wild
CC   type plants, the rice grains obtained from such transgenic plants
CC   have a favourable cooking quality (i.e. the rice grains display
CC   a reduction of water absorption) with enhanced nutritional value.
CC   The starch produced by the transgenic plants of the invention is
CC   used for a variety of purposes.
XX
SQ   Sequence    6 AA;


  Query Match             54.5%;  Score 24;  DB 21;  Length 6;
  Best Local Similarity   66.7%;  Pred. No. 6.4e+03;
  Matches    4;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;
```

```
Qy          3 KIASNTQ 9
              |:  :||:
Db          1 KSRKGT 6


RESULT   4
AAY62348
ID   AAY62348 standard; Peptide; 7 AA.
XX
AC   AAY62348;
XX
DT   02-MAR-2000  (first entry)
XX
DE   Cadherin 14 cell adhesion recognition cyclic peptide SEQ ID NO:2662.
XX
KW   Modulation; nonclassical cadherin mediated cell adhesion; CAR;
KW   inhibition; cadherin extracellular domain; cell adhesion recognition;
KW   CR cadherin; cadherin-5; cadherin 6; cadherin-7; cadherin 8;
KW   cadherin 12; cadherin-14; cadherin-15; T-cadherin; PB cadherin;
KW   cadherin related neuronal receptor; LI cadherin; protocadherin;
KW   desmerlein; desmocollin; calcium binding; cancer; tumour; leukin;
KW   rheumatoid arthritis; multiple sclerosis; diabetes; metastasis;
KW   neurological disease; cyclic.
XX
OS   Synthetic.
OS   Misc. sequences.
XX
FH   Key             Location/Qualifiers
FT   Modified site   1..7
FT                   /note= "the terminal residues are cyclized with each
FT                    other to form a cyclic peptide"
XX
PN   WO200071568-A2.
XX
PD   30-NOV-2000.
XX
PF   25 MAY 2000;   2000WO-CA00964.
XX
PR   25-MAY-1999;   99US-0073040.
PR   06-NOV-1999;   99US-0187859.
PR   20-JAN-1999;   99US-0234395.
PR   08-MAR-1999;   99US-0264516.
XX
PA   (ALBE-) ALBERTEX THERAPEUTICS INC.
XX
PI   Blaschuk OW, Gour BJ, Byers S;
XX
DR   WPI; 2000-088791/08.
XX
PT   New cadherin modulating agents, used for modulating nonclassical
PT   cadherin mediated functions for treating e.g. cancers, obesity,
PT   rheumatoid arthritis, multiple sclerosis, diabetes or a neurological
PT   disease -
XX
PS   Claim 54; Page 186; 252pp; English.
XX
CC   The present invention describes cadherin modulating agents (MA)
CC   comprising peptides which comprise a nonclassical cadherin cell adhesion
CC   recognition (CAR) sequence. The MAs can be used for modulating
CC   nonclassical cadherin-mediated functions. They can be used for e.g.
CC   inhibiting adhesion of nonclassical-cadherin expressing cells in a
CC   mammal, enhancing delivery of a drug through the skin of a mammal,
CC   enhancing delivery of a drug to a tumour in a mammal, treating cancer in
CC   a mammal, inhibiting metastasis of a cancer in a mammal, inhibiting
CC   angiogenesis in a mammal, inducing apoptosis in a nonclassical cadherin-
CC   expressing cell, preventing or treating obesity in a mammal, stimulating
CC   blood vessel regression in a mammal, enhancing drug delivery to the
CC   central nervous system, treating a demyelinating neurological disease,
CC   increasing vasopermeability in a mammal, enhancing adhesion of
CC   nonclassical cadherin expressing cells, inhibiting synaptic stability in
CC   a mammal, or preventing pregnancy in a mammal. They can also be used for
CC   e.g. enhancing or directing neurite outgrowth, facilitating wound
CC   healing or reducing scar tissue, or enhancing adhesion of foreign tissue
CC   in a mammal. They can also be used for treating e.g. psoriasis,
CC   arthritis, age related macular degeneration, multiple sclerosis and
CC   diabetes. The products can also be used for detection and diagnosis and
CC   in bioreactors. AAY60592 to AAY64772 represent specifically claimed
CC   peptides, and AAY64573 to AAY64643 and AAZ58186 to AAZ58186 represent
CC   sequences used in the exemplification of the present invention.
XX
SQ   Sequence   7 AA;


  Query Match             48.8%;  Score 21;  DB 1;  Length 7;
  Best Local Similarity   57.1%;  Pred. No. 4.6e+05;
  Matches    4;  Conservative    2;  Mismatches    1;  Indels    1;  Gaps    1;

Qy          3 KIASNTQ 9
              |:  :||:
Db          1 KIDANTG 7


RESULT   5
AAR26082
ID   AAR26082 standard; peptide; 8 AA.
XX
AC   AAR26082;
XX
DT   01-FEB-1993  (first entry)
XX
DE   Immunising peptide fragment #2 of T cell receptor
XX
KW   TCR; beta chain; rheumatoid arthritis; multiple sclerosis;
KW   autoimmune disease; diabetes; T-cell lymphoma; vaccination;
KW   immunisation.
XX
OS   Synthetic.
XX
PN   WO9212996-A.
XX
PD   06-AUG-1992.
XX
PF   21-JAN-1992;   92WO-US00482.
XX
PR   22-JAN-1991;   91US-0644611.
XX
PA   (IMMU ) IMMUNE RESPONSE CORP.
XX
PI   Brostoff SW, Carlo DJ, Howell MD;
XX
DR   WPI; 1992-295599/44.
XX
PT   Treatment of auto-immune diseases e.g. rheumatoid arthritis
PT   using vaccine comtg. T-cell receptors from surface of T cells
PI   which mediate the diseases
XX
PS   Disclosure; Page 15; 87pp; English.
XX
CC   This peptide fragment of the T cell receptor is used as an
CC   immunising peptide as part of a vaccine used to treat rheumatoid
CC   arthritis or multiple sclerosis, but can also be used against other
CC   autoimmune diseases (eg. diabetes) or T-cell lymphoma. The sequence
CC   is part of a beta chain VDJ region.
XX
SQ   Sequence   8 AA;


  Query Match             48.8%;  Score 21;  DB 1;  Length 8;
  Best Local Similarity   66.7%;  Pred. No. 5.6e+05;
  Matches    4;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          4 IASNTQ 9
              | |||:
Db          3 iqsnte 8
```

RESULT 7
AAB12940
ID AAB12940 standard; peptide; 8 AA.
XX
AC AAB12940;
XX
DT [illegible] (first entry)
XX
DE Beta chain VDJ region (human Vbeta4).
XX
KW CDR; T cell receptor; TCR; vaccine.
XX
OS Synthetic.
XX
PN WO9211359-A.
XX
PD [illegible].
XX
PF 21-DEC-1990; 90WO-US007159.
XX
PR 24-DEC-1990; 90US-0631867.
XX
PA (IMMU) IMMUNE RESPONSE CORP.
XX
PI Howell MD, Brostoff SW, Carlo DJ;
XX
DR WPI; 1992-215571/26.
XX
PT Vaccine comprising T cell receptor from T cells which mediate
PT pathology - for treating and preventing T cell lymphoma,
PT rheumatoid arthritis, autoimmune diseases etc.
XX
PS Claim 14; Page 73; 79pp; English.
XX
CC A vaccine for preventing or treating a T cell mediated pathology in
CC a vertebrate comprises a medium and a pure T cell receptor (TCR) or
CC immunogenic fragment corresp. to a TCR present on the surface of T
CC cells mediating the pathology. The immunogenic fragment may
CC comprise the amino acid sequence of a beta chain variable region,
CC pref. the Vb2 region (see features table of AAB30824 to 28).
CC Alternatively, the fragment may comprise the amino acid sequence of a
CC beta chain VDJ region of the TCR. The beta chain VDJ region may
CC comprise the sequence given in AAB42956 etc.
XX
SQ Sequence 8 AA;

Query Match 48.8%; Score 21; DB 24; Length 8;
Best Local Similarity 66.7%; Prod. No. 7.4e-05;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 4 IASNIQ 9
| :|
Db 3 IASNTE 8

RESULT 8
AAY53311
ID AAY53311 standard; Protein; 9 AA.
XX
AC AAY53311;
XX
DT 18-JAN-2000 (first entry)
XX
DE Bcr Abl epitope (aa 580 588) binds HLA B8.
XX
KW Lipopeptide; epitope; cytotoxic T lymphocyte; CTL; lipid; spacer; p53;
KW electrical charge; hydrophilicity; vaccine; immune response; HIV; HBV;
KW human immunodeficiency virus; hepatitis B virus; papilloma virus;
KW melanoma; malaria; parasite.
XX
OS Synthetic.
OS Homo sapiens.
XX
PN FR2776926-A1.
XX
PD 08-OCT-1999.
XX
PF 07-APR-1998; 98FR-0004323.
XX
PD 24-APR-2001.
XX
PF 06-JUN-1995; 95US-0471239.
XX
PR 18-JUL-1994; 94US-0276776.
PR 24-DEC-1991; 91US-0814867.
PR 21-MAR-1989; 89US-0326314.
PR 18-JUL-1989; 89US-0382086.
PR 30-MAY-1990; 90US-0530229.
PR 22-JAN-1991; 91US-0644611.
PR 18-JUL-1999; 99US-0382085.
XX
PA (IMMU) IMMUNE RESPONSE CORP.
XX
PI Howell MD, Brostoff SW, Carlo DJ;
XX
DR WPI; 2001-315571/33.
XX
PT Preventing proliferation of Vbeta14 or 17 expressing T cells, for
PT preventing, treating T cell mediated pathologies such as autoimmune
PT diseases, by administering antibody that binds to Vbeta region of T
PT cell receptor
XX
PS Disclosure; Column 6; 40pp; English.
XX
CC The specification describes a method for preventing the proliferation
CC of Vbeta14 or Vbeta17-expressing T cells in a human individual having
CC rheumatoid arthritis. The method comprises administering a cytotoxic
CC or cytostatic agent which comprises an antibody selectively binding to
CC Vbeta14 or Vbeta17 expressed by the T cells. The method is useful for
CC preventing the proliferation of T cells, which in turn is useful for
CC preventing, ameliorating or treating T cell mediated pathologies of
CC autoimmune diseases, such as rheumatoid arthritis and multiple
CC sclerosis. The present sequence represents an immunogenic peptide from
CC the B-chain VDJ region of Vbeta4 and Vbeta12, expressed by T cells
CC derived from a patient with multiple sclerosis.
XX
SQ Sequence 8 AA;

Query Match 48.8%; Score 21; DB 22; Length 8;
Best Local Similarity 66.7%; Prod. No. 7.4e-05;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 4 IASNIQ 9
| :|
Db 3 IASNTE 8

```
XX
PN   * AAY 1 ***;    98FR-000423.
XX
PA   (INRM ) INSERM INST NAT SANTE & RECH MEDICALE
PA   (CNRS ) CNRS CENT NAT RECH SCI.
PA   (INSP ) INST PASTEUR LILLE.
XX
PI   Le Gal FA,  Guillet JG,  Mabery CM,  Gras MH,  Meleys  ,  Tartar A,
XX
DR   WPI; 1999-583113/50.
XX
PT   New lipopeptide containing lipid regions and two epitopes, all
PT   separated by peptide spacers that impart hydrophilicity, useful in
PT   vaccines
XX
PS   Disclosure; Page 17  35pp; French.
XX
CC   The invention relates to the generation of a lipopeptide comprising at
CC   least one auxiliary T epitope, at least one cytotoxic T lymphocyte (CTL)
CC   epitope and at least one lipid residue with (i) the epitopes and lipid
CC   parts, and (ii) the epitopes, being separated independently by peptide
CC   spacers. These spacers comprise sequences of amino acids which carry an
CC   overall electrical charge in neutral media to ensure that the
CC   lipopeptide is hydrophilic. The peptides AAY****-AAY**** are provided
CC   examples. T peptide epitopes used to generate the lipopeptides. These are
CC   used in therapeutic or prophylactic compositions and vaccines to induce
CC   specific immune responses against human immunodeficiency, hepatitis B or
CC   papilloma viruses; p53 of melanoma or the malaria parasite.
XX
SQ   Sequence   9 AA;


  Query Match             48.8%;  Score 21;  DB 20;  Length 9;
  Best Local Similarity   66.7%;  Pred. No. 6.4e+06;
  Matches    4;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 FSKIAS 6
           | ||:|
Db       2 fqklas 7


RESULT   9
AAY40133
ID   AAY40133 standard; peptide; 9 AA.
XX
AC   AAY40133;
XX
DT   19-NOV-1999  (first entry)
XX
DE   Amino acid sequence of a chronic myeloid leukemia epitope.
XX
KW   Cytotoxic T cell; T lymphocyte; CD8+ epitope; T helper cell;
KW   CD4+ epitope; B epitope; lipopeptide; interferon gamma; IFNgamma;
KW   exercise limit; infection; immune response; cytokine profile;
KW   acquired immune deficiency syndrome; papilloma; cancer; hepatitis;
KW   autoimmune disease.
XX
OS   Homo sapiens.
XX
PN   FR2771640-A1.
XX
PD   04-JUN-1999.
XX
PF   19-FEB-1998;   98FR-0001439.
XX
PR   19-FEB-1998;   98FR-0001439.
XX
PA   (INRM ) INSERM INST NAT SANTE & RECH MEDICALE.
PA   (INSP ) INST PASTEUR LILLE.
XX
PI   Ihlam K,  Guillot JG,  Ver Waerde C,  Auriault C,  Gras MH,  Loing E;
XX
DR   WPI; 1999-312734/13.
XX
PT   New lipopeptide comprising C terminal interferon gamma fragment with
PT   attached lipophilic groups, used as interferon mimic, e.g. for treating
PT   cancer or virus infection
XX
PS   Disclosure; Page 32; 53pp; French.
XX
CC   AAY40123-AAY40379 represent epitopes that are able to activate cytotoxic
CC   T lymphocytes (CD8+ epitopes), T helper cells (CD4+ epitopes), or
CC   B epitopes recognized by corresponding antibodies. The epitopes may be
CC   used in the composition of the invention. The specification describes a
CC   lipopeptide that has a peptide part derived from mammalian interferon
CC   gamma (IFNg) and one or more lipophilic parts comprising a linear or
CC   branched, (un)saturated 4-20C hydrocarbyl chain or a steroid. The
CC   lipopeptide mimics the activity of IFNg. Compositions comprising the
CC   lipopeptide are used to treat or prevent any condition that responds
CC   to IFNg, and as adjuvant for vaccines (particularly those directed
CC   against tumors, viral or parasitic infections), to stimulate or
CC   (re)orient the immune response between Type 1 and 2 cytokine profiles.
CC   Particular applications are treatment of infections (particularly
CC   viral, e.g. acquired immune deficiency syndrome, papilloma (cancer) and
CC   hepatitis, but also bacterial, fungal, parasitic and helminth), cancers
CC   (particularly of kidney, cutaneous T cells or ovary, chronic
CC   myelogenous leukemia or mesothelioma), allergy and autoimmune
CC   diseases.
XX
SQ   Sequence   9 AA;


  Query Match             48.8%;  Score 21;  DB 20;  Length 9;
  Best Local Similarity   66.7%;  Pred. No. 6.4e+06;
  Matches    4;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 FSKIAS 6
           | ||:|
Db       2 fqklas 7


RESULT   10
AAY26649
ID   AAY26649 standard; peptide; 9 AA.
XX
AC   AAY26649;
XX
DT   14-SEP-1999  (first entry)
XX
DE   BCR ABL derived lipopeptide epitope (aa580-588) for mixed micelles.
XX
KW   Micelle; microaggregate; induction; immune response; lipopeptide; CTL;
KW   cytotoxic T lymphocyte; epitope; lipid; helper T lymphocyte; HIV; HBV;
KW   tetanus; toxin; vaccine; HIV; hepatitis B virus; papilloma virus; p53;
KW   melanoma; Plasmodium falciparum; malaria.
XX
OS   Synthetic.
OS   Abelson murine leukemia virus.
XX
PN   FR2771640-A1.
XX
PD   04-JUN-1999.
XX
PF   03-DEC-1997;   97FR-0015246.
XX
PR   03-DEC-1997;   97FR-0015246.
XX
PA   (CNRS ) CNRS CENT NAT RECH SCI.
PA   (INRM ) INSERM INST NAT SANTE & RECH MEDICALE.
PA   (INSP ) INST PASTEUR LILLE.
XX
PI   Bossus M,  Bourgault VI,  Gras-Masse H,  Guillet JG,  Lippens G;
PI   Tartar A,  Wieruszeski JM;
XX
DR   WPI; 1999-349509/30.
XX
```

```
PT   [22]Non-toxic lipopeptide micelles  comprising lipopeptide
PT   [23]Containing cytotoxic and helper T lymphocyte epitopes
XX
PS   Example; Page 60; 74pp; French.
XX
CC   The present invention relates to the generation of mixed micelles or
CC   [24]preparations for inducing an immune response comprise: (a) a first
CC   lipopeptide comprising at least one CTL (cytotoxic T lymphocyte) epitope
CC   unit at least one lipid unit; and (b) a second lipopeptide comprising at
CC   least one HTL (helper T lymphocyte) epitope and at least one lipid unit
CC   different from that of the first lipopeptide.  This peptide represents
CC   an example of a lipopeptide epitope used in the invention and is derived
CC   from the SCP protein of the Abelson murine leukaemia virus.  The
CC   non-toxic lipopeptide micelles are used in vaccines, especially against
CC   H.I.V., hepatitis B virus (HBV), papilloma viruses, p53 melanoma or
CC   falciparum malaria.
XX
SQ   Sequence  9 AA;


  Query Match             48.8%;  Score 21;  DB 20;  Length 9;
  Best Local Similarity   66.7%;  Pred. No. 6.4e+05;
  Matches    4;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       3 KIASNT 8
           :||:||
Db       2 PIKIAS 7


RESULT  1
AAW67018
ID   AAW67018 standard; Peptide; 9 AA.
XX
AC   AAW67018;
XX
DT   21-JUN-2000  (first entry)
XX
DE   Rat synaptonemal complex protein 1 (SCP1) HLA binding motif.
XX
KW   Synaptonemal complex protein 1; human; spermatocyte;
KW   Bcr-Abl; cancer; marker; breast cancer; early cancer; stroma;
KW   carcinoma; cell carcinoma; transformed cell; diagnosis; therapy;
KW   vaccine; antibody; HLA; major histocompatibility complex; MHC.
XX
OS   Rattus sp.
XX
PN   WO9934834-A1.
XX
PD   22-JUL-1999.
XX
PF   07-JAN-1999;  99WO-US000243.
XX
PR   07-JAN-1998;  98US-0070702.
XX
PA   (DUBW-) DUBWIG INST CANCER RES.
XX
IN   Fitzpatrick M;  Sahin U;  Tureci O;
XX
DR   WPI; 1999-521280/11.
XX
PT   New DNA encoding transformed cells in expression of synaptonemal complex
PT   protein - for diagnosis, prevention and treatment of cancer
XX
PS   Example 1; Page 19; 47pp; English.
XX
CC   The peptide corresponds to amino acids 415-423 of the previously
CC   known 1,234 aa synaptonemal complex protein 1 (SCP1), a protein
CC   that is restricted in meiosis and which is expressed in tumour cells,
CC   especially in renal cell carcinomas, gliomas and breast carcinomas,
CC   but not in healthy cells except for testis.  The peptide is
CC   processed to bind to MHC molecule B52.  It is one of 33 peptides (see
CC   AAW67001-033) derived from SCP 1 and identified as HLA binding motifs.
CC   Complexes formed between a peptide and an HLA molecule should
CC   provoke a cytolytic T cell response.  The invention provides
CC   immunotherapeutic approaches to conditions characterised by
CC   aberrant or abnormal levels of one or more SCP proteins.
XX
SQ   Sequence  9 AA;


  Query Match             48.8%;  Score 21;  DB 21;  Length 9;
  Best Local Similarity   66.7%;  Pred. No. 6.4e+05;
  Matches    4;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3 KIASNT 8
           :||:||
Db       2 PIKIAS 7


RESULT  12
AAY76718
ID   AAY76718 standard; Peptide; 9 AA.
XX
AC   AAY76718;
XX
DT   13-APR-2000  (first entry)
XX
DE   SCP-1 HLA binding motif #80.
XX
KW   SSX 2; SSX1 gene; breast cancer; melanoma; NY-ESO-1; SCP-1; MAGE 3;
KW   diagnosis; detection; tumour rejection antigen precursor; PCR primer;
KW   HLA binding motif.
XX
OS   Homo sapiens.
XX
PN   WO2000000642-A1.
XX
PD   06-JAN-2000.
XX
PF   03-JUN-1999;  99WO-US012293.
XX
PR   26-JUN-1998;  98US-0105907.
XX
PA   (LUDW-) LUDWIG INST CANCER RES.
XX
PI   Sahin U,  Tureci O,  Chen Y,  Old LJ,  Pfreundschuh M;
XX
DR   WPI; 2000-147070/12.
XX
PT   Detecting breast cancer and melanoma used for diagnosis and in
PT   development of targeted therapy
XX
PS   Example 2; Page 11; 24pp; English.
XX
CC   This sequence represents a HLA binding motif from the SCP-1 protein.
CC   The invention relates to a method for determining breast cancer or
CC   melanoma in a sample by assaying for expression of NY-ESO-1, SSX 2 and
CC   SCP-1 for breast cancer, or NY-ESO-1, SSX 2 and MAGE 3 for melanoma, the
CC   expression of at least one is an indicator of possible breast cancer, and
CC   the expression of at least two is an indicator of possible melanoma. The
CC   method is used to diagnose breast cancer and melanoma, and to develop
CC   therapies, e.g. a mixture of peptides, derived from tumour antigens and
CC   known to associate with major histocompatibility complex to promote
CC   cytolytic T cell proliferation, can be prepared based on the determined
CC   pattern of antigen expression in a particular sample. Detecting two or
CC   more of the specified tumour rejection antigen precursors improves the
CC   accuracy of the assay.
XX
SQ   Sequence  9 AA;


  Query Match             48.8%;  Score 21;  DB 21;  Length 9;
  Best Local Similarity   66.7%;  Pred. No. 6.4e+05;
  Matches    4;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3 KIASNT 8
```

```
          :|||:|
Db        2 qlatnd 7


RESULT  13
AAB82788
ID   AAB82788 standard; Peptide; 9 AA.
XX
AC   AAB82788;
XX
DT   29-OCT-2001  (first entry)
XX
DE   Epitope of mycobacterial antigen 85A.
XX
KW   Mycobacterial antigen 85A; Ag85A; infection; vaccine; diagnosis;
KW   therapy.
XX
OS   Mycobacterium tuberculosis.
XX
PN   WO200158481-A1.
XX
PD   16-AUG-2001.
XX
PF   12-FEB-2001; 2001WO-GB00561.
XX
PR   14-FEB-2000; 2000GB-0003082.
XX
PA   (GLAX ) GLAXO GROUP LTD.
XX
PI   [redacted] DM   Smith SM   Brookes D;
XX
DR   WPI; 2001-536505/59.
XX
PT   Use of polypeptides comprising groups of mycobacterial antigen 85A
PT   protein and expression vectors comprising polynucleotides encoding the
PT   polypeptide for vaccination against Mycobacterium infection -
XX
PS   Disclosure; Page 5; 49pp; English.
XX
CC   The present sequence is that of an epitope peptide comprising amino
CC   acid residues 242-250 of Mycobacterium tuberculosis antigen 85A
CC   (Ag85A). Epitopes, including the present sequence, have been
CC   found in Ag85A which cause the generation of a strong CD8 T-cell
CC   response in humans. The CD8 T-cells which recognise the epitopes
CC   are able to lyse macrophages infected with live Mycobacterium
CC   tuberculosis. A polypeptide which comprises an epitope structure
CC   of Ag85A, or an expression vector comprising a polynucleotide
CC   encoding such a polypeptide, is used in the manufacture of a
CC   medicament for vaccinating prophylactically or therapeutically
CC   against infection by a mycobacterium stimulating a CD8 T-cell
CC   response. The polypeptide or expression vector may be within an
CC   antigen-presenting cell. A claimed vaccine composition comprises
CC   the polypeptide or expression vector and an adjuvant or delivery
CC   system capable of stimulating a CD8 T-cell response. Also claimed
CC   are methods of detecting in a population of T-cells the presence
CC   or absence of CD8 T cells that recognise the epitope sequence, a
CC   method of diagnosing mycobacterial infection or of testing the
CC   effectiveness of a vaccination, and a method of treating a
CC   mycobacterium infection by administering T cells capable of
CC   recognising the epitope.
XX
SQ   Sequence   9 AA;


  Query Match             48.8%;  Score 21;  DB 22;  Length 9;
  Best Local Similarity   66.7%;  Pred. No. 6.4e-05;
  Matches    4;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       4 IASNIy 9
           |:|||:
Db       5 iamnir 8


RESULT  14
AAM24572
ID   AAM24572 standard; Peptide; 9 AA.
XX
AC   AAM24572;
XX
DT   04-DEC-2001  (first entry)
XX
DE   Human MHC class I molecule HLA-A1 binding 83P5G4 pept. no #4.
XX
KW   83P5G4-related protein; prostate; testis; tissue; cancer; bladder; liver;
KW   tumour; kidney; brain; bone; ovary; breast; pancreas; colon; lung; serum;
KW   cytostatic; gene therapy; antibody therapy; ribozyme; blood; cervix;
KW   single chain monoclonal antibody; urine; uterus; rectum; stomach; human;
KW   chromosome 1q31-q32.
XX
OS   Homo sapiens.
XX
PN   WO200159115-A2.
XX
PD   16-AUG-2001.
XX
PF   09-FEB-2001; 2001WO-US04426.
XX
PR   09-FEB-2000; 2000US-0181261.
XX
PA   (UROG-) UROGENESYS INC.
XX
PI   Hubert RS,  Atar DEB,  Challita-eid PM,  Faris M,  Levin E;
PI   Mitchell SC,  JaEdicke-its A;
XX
DR   WPI; 2001-514669/56.
XX
PT   An isolated 83P5G4 related protein useful as a diagnostic and/or
PT   therapeutic agent in multiple cancers such as prostate, bladder and
PT   bone cancer -
XX
PS   Example 15; Page 75; 112pp; English.
XX
CC   The polypeptide sequences represent the 83P5G4 related protein and
CC   peptide fragments of the protein. 83P5G4 exhibits prostate specific
CC   expression in normal adult tissue, but it is also aberrantly expressed in
CC   many cancers including tumours of the prostate, testis, bladder, kidney,
CC   brain, bone, cervix, uterus, ovary, breast, pancreas, stomach, rectum,
CC   liver, colon and lung. The 83P5G4 polynucleotide, its related protein and
CC   peptide fragments and specific PCR primers are therefore useful for
CC   diagnosing and treating cancer. A vector comprising a polynucleotide
CC   which encodes a single chain monoclonal antibody, that immunospecifically
CC   binds to an 83P5G4-related protein, and a ribozyme capable of cleaving a
CC   polynucleotide having the 83P5G4 coding sequence, are both useful in the
CC   preparation of a composition for treating a patient with a cancer that
CC   expresses 83P5G4. The sequences can be used in diagnostic methods to
CC   monitor the level of 83P5G4 gene products in serum, blood, urine and
CC   tissue and to thereby detect the presence of cancerous cells.
XX
SQ   Sequence   9 AA;


  Query Match             48.8%;  Score 21;  DB 22;  Length 9;
  Best Local Similarity   66.7%;  Pred. No. 6.4e-05;
  Matches    4;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 FSKIAS 6
           |:|||:
Db       4 ftkiat 9


RESULT  15
AAY61086
ID   AAY61086 standard; Peptide; 7 AA.
XX
AC   AAY61086;
XX
```

```
LI     ;  MA= 2      (first entry)
XX
TE     Modulating  cell adhesion mediated cyclic peptide SEQ ID NO:2;
XX
KW     Modulating; nonclassical cadherin mediated cell adhesion; CAM;
KW     cadherin; cadherin extracellular domain; cell adhesion; recognition;
KW     cadherin 1; cadherin 2; cadherin 3; cadherin 4; cadherin 5; cadherin 6; cadherin 7; cadherin 8;
KW     cadherin 9; cadherin 10; cadherin 11; cadherin 12; cadherin 13; cadherin 14; cadherin 15; 16 cadherin;
KW     cadherin related neuronal receptor; LI cadherin; protocadherin;
KW     desmocollin; desmoglein; calcium binding; cancer; tumour; obesity;
KW     rheumatoid arthritis; multiple sclerosis; diabetes; metastasis;
KW     neoplastic disease; cyclic.
XX
FH     Key             Location/Qualifiers
FT     Modified site   1..7
FT                     /note= "the terminal residues are released with each
FT                      other to form a cyclic peptide"
XX
PN     WO-A-01/64737.
XX
PD     13-SEP-2001.
XX
PF     07-MAR-2001;    2001WO-EP02663.
XX
PR     07-MAR-2000;    2000US-0187604P.
PR     07-NOV-2000;    2000US-0018789P.
PR     20-JAN-2001;    2001US-0244439P.
PR     08-MAR-2000;    2000US-0264516P.
XX
PA     (ALGE-) ALLERGEN TECHNOLOGIES INC.
XX
PI     Bian Hok-W;  Soan BJ;  Byers S;
XX
DR     WPI; 2001-588711/64.
XX
PT     New cadherin modulating agents, used for modulating non-classical
PT     cadherin associated functions for treating e.g. cancer, obesity,
PT     rheumatoid arthritis, multiple sclerosis, diabetes and metastasis.
XX
PS     Claim 2; Page 152pp; English.
XX
CC     The present invention describes cadherin modulating agents (MA)
CC     comprising peptides which comprise a nonclassical cadherin cell adhesion
CC     recognition (CAR) sequence. The MAs can be used for modulating
CC     non-classical cadherin mediated functions. They can be used for eg.
CC     inhibiting adhesion of nonclassical cadherin expressing cells in a
CC     mammal, enhancing delivery of a drug to the skin in a mammal,
CC     enhancing delivery of a drug to a tumour in a mammal, treating cancer in
CC     a mammal, inhibiting metastasis of a cancer in a mammal, inhibiting
CC     neoplasia in a mammal, inducing apoptosis in a nonclassical cadherin
CC     expressing cell, preventing or treating obesity in a mammal, stimulating
CC     blood vessel repression in a mammal, enhancing drug delivery to the
CC     central nervous system, treating a demyelinating neurological disease,
CC     increasing vasopermeability in a mammal, enhancing adhesion of
CC     non-classical cadherin expressing cells, inhibiting synaptic stability in
CC     a animal, or preventing pregnancy in a mammal. They can also be used for
CC     eg. enhancing or directing neurite outgrowth, facilitating wound
CC     healing, reducing scar tissue, or enhancing adhesion of foreign tissue
CC     e.g. grafts. They can also be used for treating eg. psoriasis,
CC     arthritis, age related macular degeneration, multiple sclerosis and
CC     diabetes. The products can also be used in a method of diagnosis and
CC     kits. SEQ ID NO: AAY64572 to AAY64573 represent specifically claimed
CC     SEQ ID NO: AAY64573 to AAY64618 and AAZ64584 to AAZ64586 represent
CC     peptides used in the exemplification of the present invention.
XX
SQ     Sequence  7 AA;

  Query Match          46.5%;  Score 25;  DB 21;  Length 7;
  Best Local Similarity   57.1%;  Pred. No. 6.4e-05;
  Matches    4;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       3 KIASNIQ 9
           || ||:
Db       1 kinente 7


Search completed: August 26, 2002, 06:27:10
Job time: 991 sec
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 26, 2002, 06:23:22 ; Search time 51.76 Seconds
(without alignments)
4.345 Million cell updates/sec

Title: US-09-744-282-4
Perfect score: 45
Sequence: 1 FSKIASNTQ 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: [illegible]

Minimum DB seq length: 0
Maximum DB seq length: 9

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	45	100.0	9	1	US-08-619-645-4	Sequence 4, Appli
2	45	100.0	9	2	US-08-634-433-4	Sequence 4, Appli
3	22	48.8	7	4	US-09-187-859-2062	Sequence 2062, Ap
4	22	48.8	8	2	US-08-466-860-2	Sequence 2, Appli
5	22	48.8	8	3	US-08-472-040A-2	Sequence 2, Appli
6	22	48.8	8	4	US-08-276-776-2	Sequence 2, Appli
7	22	48.8	8	4	US-08-471-209-2	Sequence 2, Appli
8	20	46.5	7	4	US-09-187-859-1032	Sequence 1032, Ap
9	20	46.5	7	4	US-09-187-859-1353	Sequence 1353, Ap
10	20	46.5	7	4	US-09-187-859-1512	Sequence 1512, Ap
11	20	46.5	7	4	US-09-187-859-2289	Sequence 2289, Ap
12	20	46.5	9	3	US-08-159-339A-1220	Sequence 1220, Ap
13	19	44.2	6	2	US-08-290-592E-47	Sequence 47, Appl
14	19	44.2	7	2	US-08-482-228-43	Sequence 43, Appl
15	19	44.2	7	3	US-08-482-528-43	Sequence 43, Appl
16	19	44.2	7	3	US-08-846-561-59	Sequence 59, Appl
17	19	44.2	7	4	US-09-187-859-2011	Sequence 2011, Ap
18	19	44.2	7	4	US-09-187-859-3265	Sequence 3265, Ap
19	19	44.2	8	2	US-08-466-860-21	Sequence 21, Appl
20	19	44.2	8	2	US-08-466-860-24	Sequence 24, Appl
21	19	44.2	8	3	US-08-472-040A-21	Sequence 21, Appl
22	19	44.2	8	3	US-08-472-040A-24	Sequence 24, Appl
23	19	44.2	8	4	US-08-276-776-21	Sequence 21, Appl
24	19	44.2	8	4	US-08-276-776-24	Sequence 24, Appl
25	19	44.2	8	4	US-08-471-209-21	Sequence 21, Appl
26	19	44.2	8	4	US-08-471-209-24	Sequence 24, Appl
27	19	44.2	8	4	US-09-187-859-2412	Sequence 2412, Ap
28	19	44.2	8	4	US-09-187-859-2663	Sequence 2663, Ap
29	19	44.2	9	1	US-07-938-[illegible]	Sequence 5, Appli
30	19	44.2	9	1	US-08-[illegible]	Sequence 18, Appl
31	19	44.2	9	1	US-08-44[illegible]	Sequence 38, Appl
32	19	44.2	9	2	US-08-490-461-22	Sequence 22, Appl
33	19	44.2	9	2	US-08-466-860-18	Sequence 18, Appl
34	19	44.2	9	3	US-08-354-626-18	Sequence 18, Appl
35	19	44.2	9	3	US-08-472-040A-28	Sequence 28, Appl
36	19	44.2	9	4	US-08-276-776-28	Sequence 28, Appl
37	19	44.2	9	4	US-08-471-209-28	Sequence 28, Appl
38	19	44.2	9	4	US-08-395-273E-28	Sequence 28, Appl
39	19	44.2	9	5	PCT-US95-04121-45	Sequence 45, Appl
40	19	44.2	9	6	5270171-4	Patent No. 5270171
41	18	41.9	7	4	US-09-187-859-1012	Sequence 1012, Ap
42	18	41.9	7	4	US-09-187-859-2333	Sequence 2333, Ap
43	18	41.9	7	4	US-09-187-859-2709	Sequence 2709, Ap
44	18	41.9	7	4	US-09-187-859-3035	Sequence 3035, Ap
45	18	41.9	8	4	US-08-444-818-641	Sequence 641, App

ALIGNMENTS

RESULT 1
US-08-619-645-4
; Sequence 4, Application US/08619645
; Patent No. 5736507
; GENERAL INFORMATION:
; APPLICANT: Boots, Anna M.H.
; APPLICANT: Verheijden, Gijsbertus F.M.
; TITLE OF INVENTION: No- 5736507al peptides derived from
; TITLE OF INVENTION: autoantigen for use in immunotherapy of ant[illegible]
; TITLE OF INVENTION: disease
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Ab- No- 5736507al Patent Department
; STREET: 1300 Piccard Drive, Suite 206
; CITY: Rockville
; STATE: Maryland
; COUNTRY: U.S.A.
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25(EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/619,645
; FILING DATE: 25-MAR-1996
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Gormley, Mary E.
; REGISTRATION NUMBER: 34,489
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 947-4433
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-619-645-4

Query Match 100.0%; Score 45; DB 1; Length 9;
Best Local Similarity 100.0%; Pred. No. [illegible];
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FSKIASNTQ 9
|||||||||
Db 1 FSKIASNTQ 9

RESULT 3
US-08-466,860-3
; Sequence 3, Application US/08466860
; Patent No. 5985552
; GENERAL INFORMATION:
; APPLICANT: HOWELL, MARK D.
; APPLICANT: BROSTOFF, STEVEN W.
; APPLICANT: CARLO, DENNIS J.
; TITLE OF INVENTION: VACCINATION AND METHODS AGAINST DISEASES
; TITLE OF INVENTION: RESULTING FROM PATHOGENIC RESPONSES BY SPECIFIC T CELL
; TITLE OF INVENTION: POPULATIONS
; NUMBER OF SEQUENCES: 75
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CAMPBELL AND FLORES
; STREET: 4370 LA JOLLA VILLAGE DRIVE, SUITE 700
; CITY: SAN DIEGO
; STATE: CALIFORNIA
; COUNTRY: UNITED STATES
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/466,860
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/813,867
; FILING DATE: 24 DEC 1991
; ATTORNEY/AGENT INFORMATION:
; NAME: CAMPBELL, CATHRYN
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-IM 9107
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-535-9001
; TELEFAX: 619-535-8949
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: N-terminal
US-08-466,860-2

Query Match 48.8%; Score 21; DB 4; Length 8;
Best Local Similarity 66.7%; Pred. No. 1.7e-05;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 3 KIASNIQ 9
|| :||:
Db 1 KIDANTE 7

RESULT 4
US-08-466,860-2
; Sequence 2, Application US/08466860
; Patent No. 5985552
; GENERAL INFORMATION:
; APPLICANT: HOWELL, MARK D.
; APPLICANT: BROSTOFF, STEVEN W.
; APPLICANT: CARLO, DENNIS J.
; TITLE OF INVENTION: VACCINATION AND METHODS AGAINST DISEASES
; TITLE OF INVENTION: RESULTING FROM PATHOGENIC RESPONSES BY SPECIFIC T CELL
; TITLE OF INVENTION: POPULATIONS
; NUMBER OF SEQUENCES: 75
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CAMPBELL AND FLORES
; STREET: 4370 LA JOLLA VILLAGE DRIVE, SUITE 700
; CITY: SAN DIEGO
; STATE: CALIFORNIA
; COUNTRY: UNITED STATES
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/466,860
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/813,867
; FILING DATE: 24 DEC 1991
; ATTORNEY/AGENT INFORMATION:
; NAME: CAMPBELL, CATHRYN
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-IM 9107
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-535-9001
; TELEFAX: 619-535-8949
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: N-terminal
US-08-466,860-2

Query Match 48.8%; Score 21; DB 4; Length 8;
Best Local Similarity 66.7%; Pred. No. 1.7e-05;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 3 KIASNIQ 9
|| :||:
Db 1 KIDANTE 7

RESULT 5
US-08-466,860-3
; Sequence 3, Application US/08466860
; Patent No. 5985552
; GENERAL INFORMATION:
; APPLICANT: HOWELL, MARK D.
; APPLICANT: BROSTOFF, STEVEN W.
; APPLICANT: CARLO, DENNIS J.
; TITLE OF INVENTION: VACCINATION AND METHODS AGAINST DISEASES
; TITLE OF INVENTION: RESULTING FROM PATHOGENIC RESPONSES BY SPECIFIC T CELL
; TITLE OF INVENTION: POPULATIONS
; NUMBER OF SEQUENCES: 75
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CAMPBELL AND FLORES
; STREET: 4370 LA JOLLA VILLAGE DRIVE, SUITE 700
; CITY: SAN DIEGO
; STATE: CALIFORNIA
; COUNTRY: UNITED STATES
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/466,860
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/813,867
; FILING DATE: 24 DEC 1991
; ATTORNEY/AGENT INFORMATION:
; NAME: CAMPBELL, CATHRYN
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-IM 9107
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-535-9001
; TELEFAX: 619-535-8949
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-466,860-3

Query Match 100.0%; Score 43; DB 4; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.7e-05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FSKIASNIQ 9
|||||||||
Db 1 FSKIASNIQ 9

RESULT 6
US-08-187,859-2062
; Sequence 2062, Application US/08187859A
; Patent No. 5585461
; GENERAL INFORMATION:
; APPLICANT: Brostoff, Steven W.
; APPLICANT: Carlo, Barbara J.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING NON-CLASSICAL
; TITLE OF INVENTION: CADHERIN MEDIATED FUNCTIONS
; NUMBER OF SEQUENCES: 2066
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Neave
; STREET: 1251 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10020-1105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/187,859A
; FILING DATE: 28-JAN-1994
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/008,407
; FILING DATE: 25-JAN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Fish, Robert E.
; REGISTRATION NUMBER: 27,983
; REFERENCE/DOCKET NUMBER: 2680/4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 596-9000
; TELEFAX: (212) 596-9090
; INFORMATION FOR SEQ ID NO: 2062:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; OTHER INFORMATION: Representative cyclic modulating agent based on
; OTHER INFORMATION: cadherin-14 cell adhesion recognition sequence
US-08-187,859-2062

Query Match 48.8%; Score 21; DB 4; Length 7;
Best Local Similarity 57.1%; Pred. No. 1.7e-05;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 3 KIASNIQ 9
|| :||:
Db 1 KIDANTE 7

RESULT 7
US-08-187,859-2062
; Sequence 2062, Application US/08187859A
; Patent No. 5585461
; GENERAL INFORMATION:
; APPLICANT: Brostoff, Steven W.
; APPLICANT: Carlo, Barbara J.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING NON-CLASSICAL
; TITLE OF INVENTION: CADHERIN MEDIATED FUNCTIONS
; NUMBER OF SEQUENCES: 2066
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Neave
; STREET: 1251 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10020-1105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/187,859A
; FILING DATE: 28-JAN-1994
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/008,407
; FILING DATE: 25-JAN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Fish, Robert E.
; REGISTRATION NUMBER: 27,983
; REFERENCE/DOCKET NUMBER: 2680/4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 596-9000
; TELEFAX: (212) 596-9090
; INFORMATION FOR SEQ ID NO: 2062:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; OTHER INFORMATION: Representative cyclic modulating agent based on
; OTHER INFORMATION: cadherin-14 cell adhesion recognition sequence
US-08-187,859-2062

Query Match 48.8%; Score 21; DB 4; Length 7;
Best Local Similarity 66.7%; Pred. No. 1.7e-05;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 3 KIASNIQ 9
|| :||:
Db 1 KIDANTE 7

```
Qy       4 IASNIQ 9
           I III:
Db       3 IGSNIE 8


RESULT   5
US-08-472-040A-2
; Sequence 2, Application US/08472040A
; Patent No. 6096687
;  GENERAL INFORMATION:
;    APPLICANT:  HOWELL, MARK D.
;    APPLICANT:  BROSNIFF, STEVEN W.
;    APPLICANT:  CARLO, DENNIS J.
;    TITLE OF INVENTION:  VACCINATION AND METHODS AGAINST DISEASES
;    TITLE OF INVENTION:  RESULTING FROM PATHOGENIC RESPONSES BY SPECIFIC T CELL
;    TITLE OF INVENTION:  POPULATIONS
;    NUMBER OF SEQUENCES:  77
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  CAMPBELL & FLORES LLP
;      STREET:  4370 LA JOLLA VILLAGE DRIVE, SUITE 700
;      CITY:  SAN DIEGO
;      STATE:  CALIFORNIA
;      COUNTRY:  UNITED STATES
;      ZIP:  92122
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/472,040A
;      FILING DATE:  06-JUN-1995
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  CAMPBELL, CATHRYN
;      REGISTRATION NUMBER:  41,815
;      REFERENCE/DOCKET NUMBER:  P-IM 1641
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  619-535-9001
;      TELEFAX:  619-535-8949
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  8 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    FRAGMENT TYPE:  N-terminal
US-08-472-040A-2


  Query Match             48.8%;  Score 21;  DB 4;  Length 8;
  Best Local Similarity   66.7%;  Pred. No. 1.7e+05;
  Matches    4;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       4 IASNIQ 9
           I III:
Db       3 IGSNIE 8


RESULT   6
US-08-276-776-2
; Sequence 2, Application US/08276776
; Patent No. 6207645
;  GENERAL INFORMATION:
;    APPLICANT:  HOWELL, MARK D.
;    APPLICANT:  BROSNIFF, STEVEN W.
;    APPLICANT:  CARLO, DENNIS J.
;    TITLE OF INVENTION:  VACCINATION AND METHODS AGAINST DISEASES
;    TITLE OF INVENTION:  RESULTING FROM PATHOGENIC RESPONSES BY SPECIFIC T CELL
;    TITLE OF INVENTION:  POPULATIONS
;    NUMBER OF SEQUENCES:  75
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  CAMPBELL AND FLORES
;      STREET:  4370 LA JOLLA VILLAGE DRIVE, SUITE 700
;      CITY:  SAN DIEGO
;      STATE:  CALIFORNIA
;      COUNTRY:  UNITED STATES
;      ZIP:  92122
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/276,776
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/813,867
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  CAMPBELL, CATHRYN
;      REGISTRATION NUMBER:  41,815
;      REFERENCE/DOCKET NUMBER:  P-IM 9167
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  619-535-9001
;      TELEFAX:  619-535-8949
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  8 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    FRAGMENT TYPE:  N-terminal
US-08-276-776-2


  Query Match             48.8%;  Score 21;  DB 4;  Length 8;
  Best Local Similarity   66.7%;  Pred. No. 1.7e+05;
  Matches    4;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       4 IASNIQ 9
           I III:
Db       3 IGSNIE 8


RESULT   7
US-08-471-209-2
; Sequence 2, Application US/08471209
; Patent No. 6221352
;  GENERAL INFORMATION:
;    APPLICANT:  HOWELL, MARK D.
;    APPLICANT:  BROSNIFF, STEVEN W.
;    APPLICANT:  CARLO, DENNIS J.
;    TITLE OF INVENTION:  VACCINATION AND METHODS AGAINST DISEASES
;    TITLE OF INVENTION:  RESULTING FROM PATHOGENIC RESPONSES BY SPECIFIC T CELL
;    TITLE OF INVENTION:  POPULATIONS
;    NUMBER OF SEQUENCES:  75
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  CAMPBELL AND FLORES
;      STREET:  4370 LA JOLLA VILLAGE DRIVE, SUITE 700
;      CITY:  SAN DIEGO
;      STATE:  CALIFORNIA
;      COUNTRY:  UNITED STATES
;      ZIP:  92122
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/471,209
;      FILING DATE:
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/813,867
```

```
;       FILING DATE:  24-DEC-1991
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  CAMPBELL, CATHRYN
;       REGISTRATION NUMBER:  31,815
;       REFERENCE/DOCKET NUMBER:  P-IM-91107
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  619-535-9001
;       TELEFAX:  619-535-8949
;   INFORMATION FOR SEQ ID NO:  2:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  8 amino acids
;       TYPE:  amino acid
;       STRANDEDNESS:  linear
;     MOLECULE TYPE:  peptide
;     FRAGMENT TYPE:  N-terminal
US-08-471,209-2

  Query Match             46.8%;  Score 21;  DB 4;  Length 8;
  Best Local Similarity   66.7%;  Prod. No. 1.7e+05;
  Matches    4;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       3 KIASNTQ 9
           |  |  |
Db       1 KIDENIE 7


RESULT  9
US-09-187-859-1512
; Sequence 1512, Application US/09187859A
; Patent No. 6358920
; GENERAL INFORMATION:
;  APPLICANT: Blaschuk, Orest W.
;  APPLICANT: Gour, Barbara J.
;  TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING N-CADHERIN
;  TITLE OF INVENTION: CADHERIN MEDIATED FUNCTIONS
;  FILE REFERENCE: 100086.407C1
;  CURRENT APPLICATION NUMBER: US/09/187,859A
;  CURRENT FILING DATE: 1998-11-06
;  NUMBER OF SEQ ID NOS: 4052
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1512
;   LENGTH: 7
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Representative cyclic modulating agent based on
;   OTHER INFORMATION: cadherin-7 cell adhesion recognition sequence
US-09-187-859-1512

  Query Match             46.8%;  Score 21;  DB 4;  Length 7;
  Best Local Similarity   66.7%;  Prod. No. 1.7e+05;
  Matches    4;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       3 KIASNTQ 9
           |  |  |
Db       1 KIDENIE 7


RESULT  10
US-09-187-859-1353
; Sequence 1353, Application US/09187859A
; Patent No. 6358920
; GENERAL INFORMATION:
;  APPLICANT: Blaschuk, Orest W.
;  APPLICANT: Gour, Barbara J.
;  TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING N-CADHERIN
;  TITLE OF INVENTION: CADHERIN MEDIATED FUNCTIONS
;  FILE REFERENCE: 100086.407C1
;  CURRENT APPLICATION NUMBER: US/09/187,859A
;  CURRENT FILING DATE: 1998-11-06
;  NUMBER OF SEQ ID NOS: 4052
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1353
;   LENGTH: 7
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Representative cyclic modulating agent based on
;   OTHER INFORMATION: cadherin-7 cell adhesion recognition sequence
US-09-187-859-1353

  Query Match             46.8%;  Score 20;  DB 4;  Length 7;
  Best Local Similarity   57.1%;  Prod. No. 1.7e+05;
  Matches    4;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       3 KIASNTQ 9
           |  |  |
Db       1 KIDENIE 7


RESULT  11
US-09-187-859-2289
; Sequence 2289, Application US/09187859A
; Patent No. 6358920
; GENERAL INFORMATION:
;  APPLICANT: Blaschuk, Orest W.
;  APPLICANT: Gour, Barbara J.
;  TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING N-CADHERIN
;  TITLE OF INVENTION: CADHERIN-MEDIATED FUNCTIONS
;  FILE REFERENCE: 100086.407C1
;  CURRENT APPLICATION NUMBER: US/09/187,859A
;  CURRENT FILING DATE: 1998-11-06
;  NUMBER OF SEQ ID NOS: 4052
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2289
;   LENGTH: 7
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Representative cyclic modulating agent based on
```

```
;  OTHER INFORMATION: E-cadherin cell adhesion recognition sequence
US-08-282-854-2289


  Query Match              46.5%;  Score 20;  DB 4;  Length 7;
  Best Local Similarity    57.1%;  Pred. No. 1.7e+06;
  Matches    4;  Conservative    1;  Mismatches    2;  Indels    1;  Gaps    1;

Qy       6 KIASNIQ 12
           |:|  |
Db       1 KINENIE 7


RESULT  13
US-08-290-592E-46
; Sequence 46, Application US/08290592E
; Patent No. 5824307
;  GENERAL INFORMATION:
;    APPLICANT:  JOHNSON, L.
;    TITLE OF INVENTION:  Human Murine Chimeric Antibodies Fragments
;    TITLE OF INVENTION:  Respiratory Syncytial Virus
;    NUMBER OF SEQUENCES:  49
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  CARELLA, BYRNE, BAIN, GILFILLAN, CECCHI, STEWART &
;      ADDRESSEE:  OLSTEIN
;      STREET:  6 BECKER FARM ROAD
;      CITY:  ROSELAND
;      STATE:  NEW JERSEY
;      COUNTRY:  USA
;      ZIP:  07068
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5 INCH DISKETTE
;      COMPUTER:  IBM PS/2
;      OPERATING SYSTEM:  MS-DOS
;      SOFTWARE:  Word Perfect 5.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/290,592E
;      FILING DATE:  August 15, 1994
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/813,372
;      FILING DATE:  December 23, 1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Olstein, Elliot M.
;      REGISTRATION NUMBER:  24,025
;      REFERENCE/DOCKET NUMBER:  469201.257
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  201 994-1700
;      TELEFAX:  201-994-1744
;  INFORMATION FOR SEQ ID NO:  47:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  6 AMINO ACIDS
;      TYPE:  AMINO ACID
;      TOPOLOGY:  LINEAR
;    MOLECULE TYPE:  PROTEIN
US-08-290-592E-47


  Query Match              44.2%;  Score 19;  DB 2;  Length 7;
  Best Local Similarity    80.0%;  Pred. No. 1.7e+05;
  Matches    4;  Conservative    1;  Mismatches    0;  Indels    1;  Gaps    0;

Qy       2 SKIAS 6
           |:|||
Db       2 SKLAS 6


RESULT  14
US-08-482-228-48
; Sequence 48, Application US/08482228
; Patent No. 5968753
;  GENERAL INFORMATION:
;    APPLICANT:  Isemi Law, Janet
;    APPLICANT:  Koberi, Joan A.
;    APPLICANT:  Al Abdaly, Fahad A.
;    APPLICANT:  Guillermo, Roy
;    APPLICANT:  Bolgerson, Sam L.
;    APPLICANT:  Deans, Robert J.
;    TITLE OF INVENTION:  POSITIVE AND NEGATIVE CELL
;    TITLE OF INVENTION:  SELECTION MEDIATED BY PEPTIDE RELEASE
;    NUMBER OF SEQUENCES:  215
;    CORRESPONDENCE ADDRESSEE:
;      ADDRESSEE:  Janice Guthrie, Ph.D.
;      STREET:  P.O. Box 15210
```

; CITY: Irvine
; STATE: California
; COUNTRY: USA
; ZIP: 92715-5210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC DOS/MS DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/482,528
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Guthrie, Janice
; REGISTRATION NUMBER: 35,170
; REFERENCE/DOCKET NUMBER: IT 4630CIP3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (714) 440-5353
; TELEFAX: (714) 553-1952
; INFORMATION FOR SEQ ID NO: 43:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-482-528-43

Query Match 44.2%; Score 19; DB 3; Length 7;
Best Local Similarity 80.0%; Pred. No. 1.7e+05;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 SKIAS 6
||:||
Db 3 SKLAS 7

RESULT 2
US-08-482-528-43
; Sequence 43, Application US/08482528
; Patent No. 05627179
; GENERAL INFORMATION:
; APPLICANT: Isobel Law, Janet
; APPLICANT: Robert, Joan A.
; APPLICANT: Al Abdaly, Fahad A.
; APPLICANT: Guillermo, Roy
; APPLICANT: Bellerose, Sam S.
; APPLICANT: Deane, Robert J.
; TITLE OF INVENTION: POSITIVE AND POSITIVE/NEGATIVE SELE
; TITLE OF INVENTION: SELECTION MEDIATED BY PEPTIDE RELEASE
; NUMBER OF SEQUENCES: 275
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Janice Guthrie, Ph.D.
; STREET: P.O. Box 19215
; CITY: Irvine
; STATE: California
; COUNTRY: USA
; ZIP: 92713-5210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC DOS/MS DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/482,528
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Guthrie, Janice
; REGISTRATION NUMBER: 35,170
; REFERENCE/DOCKET NUMBER: IT 4630CIP3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (714) 440-5353
; TELEFAX: (714) 553-1952
; INFORMATION FOR SEQ ID NO: 43:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-482-528-43

Query Match 44.2%; Score 19; DB 3; Length 7;
Best Local Similarity 80.0%; Pred. No. 1.7e+05;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 SKIAS 6
||:||
Db 3 SKLAS 7

Search completed: August 26, 2002, 06:23:22
Job time: 763 sec

SeqCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

Main Test: Protein Search, using sw model

Run On: August 26, 2002, 06:11:33 ; Search Time 5.25 Seconds
(without alignments)
24.760 Million cell updates/sec

Title: US-09-744-282-5
Perfect score: 24
Sequence: 1 PTPEGKSFTLASSE 14

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 1824

Minimum DB seq length: 0
Maximum DB seq length: 14

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 71:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	24	96.9	11	2	C58501	42K bile stone pro
2	24	96.9	13	2	PH0805	T-cell receptor al
3	23	92.3	9	2	PI0324	Ig heavy chain CRD
4	23	92.3	10	2	I28027	protein 27 - cycle
5	22	88.8	11	2	A40693	transgelin - sheep
6	20	80.8	11	2	PT0250	Ig heavy chain CRD
7	20	80.8	12	2	I46922	gene Beta protein
8	20	80.8	12	2	F20907	Ig kappa 2 chain J
9	20	80.8	12	2	PH1457	T-cell receptor be
10	20	80.8	13	1	NTKNA3	alpha-endosulfa 31
11	20	80.8	13	2	PQ0491	Sec1 incompatibili
12	20	80.8	13	2	D56661	S. cerevisiae specific a
13	19	76.2	8	2	S21273	cellulase (EC 3.2.
14	19	76.2	12	2	PH1462	T-cell receptor be
15	19	76.2	13	1	NTKN2C	alpha-endosulfa 61
16	19	76.2	13	2	A32486	beta protein - rat
17	18	73.7	4	2	S53608	glycerol-3-phosphat
18	18	73.7	8	2	A37521	R-phycoerythrin ga
19	18	73.7	10	2	S38305	rectin GNL2 alpha
20	18	73.7	10	2	A27617	transaldolase phosphate i
21	18	73.7	10	2	B61440	pol5 antarctic-phase
22	18	73.7	10	2	S65387	cytochrome c oxida
23	18	73.7	11	2	A57458	gene Gax protein -
24	18	73.7	12	2	PH1461	T-cell receptor be
25	18	73.7	13	2	S23638	Ig kappa chain J s
26	18	73.7	13	2	E53275	Ig kappa-1 chain I
27	18	73.7	13	2	PH0756	T-cell receptor be
28	17.5	26.9	8	2	A32523	peptidyl dipeptida
29	17.5	26.9	13	2	PC2471	proteoglycan endopepti
30	17	26.2	5	2	PH0615	T-cell receptor be
31	17	26.2	5	2	PI1535	T-cell receptor be
32	17	26.2	5	2	PI0464	T-cell receptor be
33	17	26.2	6	2	PI3048	T-cell receptor be
34	17	26.2	7	2	A25269	ras guanosine GAP
35	17	26.2	7	2	PH0529	T-cell receptor be
36	17	26.2	7	2	PI0607	T-cell receptor be
37	17	26.2	7	2	PI3702	T-cell receptor be
38	17	26.2	8	2	PH0803	T-cell receptor al
39	17	26.2	8	2	PT0595	T-cell receptor be
40	17	26.2	9	2	S59265	small acidic prot
41	17	26.2	9	2	B38243	Ig kappa chain V-I
42	17	26.2	9	2	PB0935	T-cell receptor be
43	17	26.2	9	2	PH0917	T-cell receptor be
44	17	26.2	9	2	PH0923	T-cell receptor be
45	17	26.2	11	2	A44795	Golgin by homolo

ALIGNMENTS

RESULT 1
C58501
42K bile stone protein - unidentified bacterium (fragment)
C;Species: unidentified bacterium
C;Date: 07-Feb-1997 #sequence_revision 07-Feb-1997 #text_change 30-Jul 2008
C;Accession: C58501
R;Binette, J.P.; Binette, M.B.
submitted to the Protein Sequence Database, October 1996
A;Description: The proteins of kidney and gallbladder stones.
A;Reference number: A58501
A;Accession: C58501
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-11 <BIN>
A;Experimental source: human bile with stones
A;Note: tentitive identification of 1-31y

Query Match 96.9%; Score 24; DB 2; Length 11;
Best Local Similarity 100.0%; Prod. No. 2.26e-2;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PTPEG 4
||||
Db 6 PTEG 9

RESULT 2
PH0805
T-cell receptor alpha chain (PES.1.IV alpha 2.8) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 09-May-1997
C;Accession: PH0805
R;Casanova, J.L.; Romero, P.; Widmann, C.; Kourilsky, P.; Maryanski, J.L.
J. Exp. Med. 174, 1421-1434, 1991
A;Title: T cell receptor genes in a series of class I major histocompatibility complex allelic exclusion and antigen-specific repertoire.
A;Reference number: PH0740; MUID:92078846
A;Accession: PH0805
A;Molecule type: mRNA
A;Residues: 1-13 <CAS>
A;Cross-references: EMBL:X66914
A;Experimental source: T lymphocyte
C;Keywords: T-cell receptor

Query Match 96.9%; Score 24; DB 2; Length 13;
Best Local Similarity 100.0%; Prod. No. 2.66e-2;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PTPEG 4
||||

Qy 3 PIEGRS 8

RESULT 4
PI0234
Ig heavy chain CRO4 region (clone J2-1067) - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 30-Sep-1993 #Sequence_revision 30-Sep-1993 #text_change 16-Aug-1996
C;Accession: PI0234
R;Yamada, M.; Wasserman, R.; Reichard, B.A.; Shane, S.; Caton, A.J.; Rovera, G.
J. Exp. Med. 173, 395-407, 1991
A;Title: Preferential utilization of specific immunoglobulin heavy chain diversity and j
A;Reference number: PI0222; MUID:91108437
A;Accession: PI0234
A;Molecule type: DNA
A;Residues: 1-9 <YAM>
A;Experimental source: B lymphocyte
C;Keywords: heterotetramer; immunoglobulin

Query Match 42.8%; Score 21; DB 2; Length 9;
Best Local Similarity 42.9%; Pred. No. 2.8e+05;
Matches 3; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 2 IPEGRSF 7
Db 3 VPSGMS 8

RESULT 5
A36927
leaf protein - tobacco (fragment)
C;Species: Nicotiana plumbaginifolia (curled leaved tobacco)
C;Date: 29-May-1989 #Sequence_revision 29-May-1989 #text_change 19-Jun-1998
C;Accession: A36927
R;Pana, J.-P.; de Looose, M.; Inze, D.; Van Montagu, M.; Vandekerckhove, J.
Proc. Natl. Acad. Sci. U.S.A. 84, 4896-4901, 1987
A;Title: Alterations in the phenotype of plant cells studied by NH2 terminal amino acid
A;Reference number: A36167
A;Accession: A36927
A;Molecule type: protein
A;Residues: 1-12 <BAU>

Query Match 42.8%; Score 21; DB 2; Length 12;
Best Local Similarity 100.0%; Pred. No. 7.2e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 PEG 5
Db 9 PEG 11

RESULT 6
A40603
tropomyosin - sheep (fragment)
C;Species: Ovis aries, Ovis ammon aries (domestic sheep)
C;Date: 1-May-1994 #Sequence_revision 03-May-1994 #text_change 31-Oct-1997
C;Accession: A40603
R;Shapland, C.; Hsuan, J.J.; Totty, N.F.; Lawson, D.
J. Cell Biol. 121, 1065-1074, 1993
A;Title: Purification and properties of transgelin: a transformation and shape change su
A;Reference number: A40603; MUID:93273760
A;Accession: A40603
A;Molecule type: protein
A;Residues: 1-11 <SHA>
A;Experimental source: aorta
C;Comment: This protein binds actin and is shown reversibly translocated or lost of co
C;Superfamily: smooth muscle protein SM22; calponin repeat homology; smooth muscle pr
C;Keywords: actin binding; cytoskeleton

Query Match 40.8%; Score 20; DB 2; Length 11;
Best Local Similarity 50.0%; Pred. No. 1.6e+05;
Matches 3; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PIEGRS 6
Db 3 PSYGMS 8

RESULT 6
PT0250
Ig heavy chain CRO3 region (clone 2-109B) - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 30-Sep-1993 #Sequence_revision 30-Sep-1993 #text_change 16-Aug-1996
C;Accession: PT0250
R;Yamada, M.; Wasserman, R.; Reichard, B.A.; Shane, S.; Caton, A.J.; Rovera, G.
J. Exp. Med. 173, 395-407, 1991
A;Title: Preferential utilization of specific immunoglobulin heavy chain diversity an
A;Reference number: PT0222; MUID:91108437
A;Accession: PT0250
A;Molecule type: DNA
A;Residues: 1-11 <YAM>
A;Experimental source: B lymphocyte
C;Keywords: heterotetramer; immunoglobulin

Query Match 30.8%; Score 20; DB 2; Length 11;
Best Local Similarity 50.0%; Pred. No. 1.6e+05;
Matches 3; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 3 PGRSFI 8
Db 6 PGRPYS 11

RESULT 7
I46922
gene Bota protein - bovine (fragment)
C;Species: Bos primigenius taurus (cattle)
C;Date: 21-Feb-1997 #Sequence_revision 21-Feb-1997 #text_change 07-Nov-1999
C;Accession: I46922
R;Ellis, S.A.; Braem, K.A.; Morrison, W.I.
Immunogenetics 37, 49-56, 1992
A;Title: Transmembrane and cytoplasmic domain sequences demonstrate at least two expr
A;Reference number: I46921; MUID:93052564
A;Accession: I46922
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-12 <ELL>
A;Cross-references: GB:S47738; NID:g258999; PIDN:AAB23973.1; PID:g258999
C;Genetics:
A;Gene: Bota

Query Match 30.8%; Score 20; DB 2; Length 12;
Best Local Similarity 50.0%; Pred. No. 1.6e+05;
Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 4 GRSFTLAS 11
Db 4 GKIYIQAA 11

RESULT 8
F20907
Ig kappa 2 chain J5 chain - rabbit
C;Species: Oryctolagus cuniculus (domestic rabbit)
C;Date: 10-Aug-1990 #Sequence_revision 10-Aug-1990 #text_change 09-Nov-1990
C;Accession: F20907
R;Emorine, L.; Max, E.E.
Nucleic Acids Res. 11, 8877-8890, 1983
A;Title: Structural analysis of a rabbit immunoglobulin kappa J-C locus reveals mult
A;Reference number: A20907; MUID:84169523
A;Accession: F20907

A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1 [illegible] <EM>
A;Cross-reference: GB:X[illegible]; NID:g1582; PIDN:CAA2[illegible]; PID:g[illegible]
C;Keywords: heterotetramer; immunoglobulin

Query Match 30.8%; Score 20; DB 2; Length [illegible];
Best Local Similarity 50.0%; Pred. No. 1.4e-03;
Matches 4; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 2 IFSSFI 9
Db 2 IFGKSKL 9

RESULT 9
PH1457
T-cell receptor beta chain (clone A5/D11) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 10-Mar-1994 #sequence_revision 10-Mar-1994 #text_change 17-Mar-1995
C;Accession: PH1457; S26542
R;Casanova, J.L.; Martinez, F.; Gournier, H.; Barra, C.; Fournier, C.; Bonnaull, A.; Ko
J. Exp. Med. 177, 811-820, 1993
A;Title: T cell receptor selection by and recognition of two class I major histocompatib
A;Reference number: PH1430; MUID:93171821
A;Accession: PH1457
A;Molecule type: mRNA
A;Residues: 1-12 <VAS>
A;Experimental source: cytolytic T lymphocyte, clone A5/D11
R;Casanova, J.L.; Cerottini, J.C.; Matthes, M.; Necker, A.; Widmer, H.; Barra, C.; Wil
J. Exp. Med. 175, 439-447, 1992
A;Title: H-2 restricted cytolytic T lymphocytes specific for E1A display T cell receptor
A;Reference number: S26512; MUID:92364546
A;Accession: S26542
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-12 <VAS>
A;Cross-reference: EMBL:X67992
A;Experimental source: cytolytic T lymphocyte, clone A5/D11
C;Superfamily: immunoglobulin homology
C;Keywords: receptor; T cell

Query Match 30.8%; Score 20; DB 2; Length 12;
Best Local Similarity 42.9%; Pred. No. 1.4e-03;
Matches 3; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 2 IFSSFI 8
Db 4 SFSSYI 10

RESULT 10
N1KNAS
alpha-cono*in SIA - cone shell (Conus striatus)
C;Species: Conus striatus (striated cone)
C;Date: 30-Sep-1992 #sequence_revision 30-Sep-1992 #text_change 25-May-1995
C;Accession: A40312
R;Myers, R.A.; Zafaralla, G.C.; Gray, W.R.; Abbott, J.; Cruz, L.J.; Olivera, B.M.
Biochemistry 30, 9370-9377, 1991
A;Title: alpha-Conotoxins, small peptide probes of nicotinic acetylcholine receptors.
A;Reference number: A40312; MUID:91369955
A;Accession: A40312
A;Molecule type: protein
A;Residues: 1-15 <MYE>
C;Comment: This paralytic toxin from a fish hunting cone snail inhibits the acetylchol
C;Superfamily: alpha-conotoxin
C;Keywords: acetylcholine receptor inhibitor; amidated carboxyl end; postsynaptic neuro
F;2-7,3-13/Disulfide bonds: #status experimental
F;15/Modified site: amidated carboxyl end (Cys) #status experimental

Query Match 30.8%; Score 20; DB 2; Length 15;
Best Local Similarity 42.9%; Pred. No. [illegible];
Matches 3; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 PIFSSRSF 7
Db 5 PAVGKNF 11

RESULT 11
PQ0491
self incompatibility locus glycoprotein delta - wild cabbage (fragment)
C;Species: Brassica oleracea (wild cabbage)
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 31-Oct-1997
C;Accession: PQ0491
R;Gaude, T.; Friry, A.; Heizmann, P.; Mariac, C.; Rougier, M.; Fobis, I.; Dumas, C.
Plant Cell 5, 75-86, 1993
A;Title: Expression of a self-incompatibility gene in a self compatible line of Brass
A;Reference number: JQ1733; MUID:93177215
A;Accession: PQ0491
A;Molecule type: protein
A;Residues: 1-13 <GAU>
A;Experimental source: stigma, var. acephala 85?si
C;Superfamily: S-receptor kinase; protein kinase homology; S locus specific glycoprot
C;Keywords: glycoprotein

Query Match 30.8%; Score 20; DB 2; Length 13;
Best Local Similarity 66.7%; Pred. No. 1.4e-03;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 8 TLASSE 13
Db 3 TLSSTE 8

RESULT 12
D56661
S-locus specific glycoprotein (allele S8) - wild cabbage (fragment)
C;Species: Brassica oleracea (wild cabbage)
C;Date: 05-Jan-1996 #sequence_revision 05-Jan-1996 #text_change 05-Jan-1996
C;Accession: D56661
R;Gaude, T.; Denoroy, L.; Dumas, C.
Electrophoresis 12, 646-654, 1991
A;Title: Use of a fast protein electroblotter purification and N-terminal
A;Reference number: A56661; MUID:92070697
A;Accession: D56661
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-13 <GAU>
A;Experimental source: stigma extracts, var. acephala
A;Note: sequence extracted from NCBI backbone (NCBIP:[illegible])
C;Comment: This glycoprotein, expressed only in stigmas, plays an important role in t
C;Keywords: glycoprotein; polymorphism

Query Match 30.8%; Score 20; DB 2; Length 15;
Best Local Similarity 66.7%; Pred. No. [illegible];
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 8 TLASSE 13
Db 3 TLSSTE 8

RESULT 13
S21273
cellulase (EC 3.2.1.4) - Clostridium thermocellum (fragment)
N;Alternate names: endo-1,4-beta-glucanase
C;Species: Clostridium thermocellum
C;Date: 04-Dec-1992 #sequence_revision 04-Dec-1992 #text_change 22-Nov-1996
C;Accession: S21273
R;Romaniec, M.P.M.; Fauth, U.; Kobayashi, T.; Buckland, N.J.; Barker, I.; Demain,

```
Publication:  J. Biol. Chem., 1992
Article:  Isolation and characterization of a new cDNA...  mouse from Destridium therm
Author(s):  Kuhara, T.; Inoue, Y.; Matsumoto, M.; Inoue, I.
Reference Number:  A26273; MUID:92243858
Accession:  A26273
Molecule type:  protein
Residues:  1-496 <KUH>
Cross-reference:
Description:  cDNA; hydrolysis of 1,4 beta D-glucosidic linkages in beta-glucans such as c
Function:  cellulose degradation
Keywords:  glycosidase; hydrolase; polysaccharide degradation


  Query Match             29.2%;  Score 19;  DB 2;  Length 496;
  Best Local Similarity   50.0%;  Prod; No. 2;  Conservative    1;  Mismatches    2;  Indels    1;  Gaps    0;

Qy          1 MIF 3

Db        496 MIF 498


Result    3
ID  Ig
Ig chain receptor beta chain (clone A24/PEF1)  mouse (fragment)
Species:  Mus musculus (house mouse)
Date:  Mar-1991 #Sequence revision 10-Mar-1991 #Text change 22-Apr-1999
Accession:  S01402
Authors:  Kappler, J.; Martinson, J.; Gounter, B.; Burtis, K.; Freed...; Kotzauli, A.; Ko
Title:  Exp. Med., 173, 1427-1432, 1991
Journal:  Nature, 322, 145-149, 1986
Abstract:  T-cell receptor selection by and recognition of the class I major histocompatib
Reference Number:  S01402; MUID:91351821
Accession:  S01402
Molecule type:  mRNA
Residues:  1-312 <KAP>
Cross-reference:
Experimental source:  cytolytic T lymphocyte
Superfamily:  immunoglobulin homology
Keywords:  receptor; T-cell


  Query Match             29.2%;  Score 19;  DB 2;  Length 312;
  Best Local Similarity   42.9%;  Prod; No. 2;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 MIFGRSET 8
                :
Db        312 MIFGKSET 319


Result    4
NTEN.1
alpha-1-conotoxin GI  cone shell (Conus geographus)
Species:  Conus geographus (geographer cone)
Date:  14-Sep-1981 #Sequence revision 24-Sep-1981 #Text change 24-May-1987
Accession:  A01583
Authors:  Gray, W.R.; Luque, A.; Olivera, B.M.; Barrett, J.; Cruz, L.J.
Citation:  Proc. Natl. Acad. Sci. U.S.A. 78, 4734-4738, 1981
Title:  Peptide toxins from Conus geographus venom.
Reference Number:  A02420; MUID:81191854
Accession:  A01583
Molecule type:  protein
Residues:  1-18 <GRA>
Authors:  Gray, W.R.; Olivera, B.M.; Cruz, L.J.  et al. B.M.
Cross-reference:  PROC. Natl. Acad. Sci. U.S.A. 78, 4734-4738, 1981
Title:  alpha-conotoxin GI, disulfide bonding and conformational states.
Reference Number:  A02420; MUID:84043400
Accession:  disulfide bonds
Comment:  Its alpha conotoxin, as an acetylcholine receptor inhibitor, is a postsynapt
Superfamily:  alpha conotoxin
Keywords:  acetylcholine receptor inhibitor; blocked carboxy end; postsynaptic neurote
Features:  disulfide bonds: #status predicted
Fold/Modified:  blocked carboxyl end (Cys) (partially amidated) #status experimental


  Query Match             29.2%;  Score 19;  DB 1;  Length 18;
  Best Local Similarity   47.6%;  Prod; No. 2;  Conservative    2;  Mismatches    3;  Indels    1;  Gaps    1;

Qy          1 MIFGRSET 8
                | | | |
Db          5 IACGKHFS 12


Search completed:  August 26, 2002, 06:12:49
Job time:  76 sec
```

GeneCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 26, 2002, 06:13:48 ; Search time 20.5 Seconds
(without alignments)
18.936 Million cell updates/sec

Title: US-09-744-282-5
Perfect score: 65
Sequence: 1 PFFGRSFILASSE 13

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0
Maximum DB seq length: 13

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt 40:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	29	44.6	12	1	V25K_WSSV	P82004 white spot
2	28	43.1	13	1	CXAA_CONST	P28878 conus stria
3	19	29.2	9	1	FAR5_ASCSU	P43170 ascaris sum
4	19	29.2	13	1	CXA2_CONGE	P0C7.. conus geogr
5	18	27.7	9	1	FARP_CALST	P38419 callinectes
6	18	27.7	10	1	AKHX_LOCMI	P49.. locusta mig
7	18	27.7	10	1	COXM_RAT	P0C435 rattus norv
8	18	27.7	10	1	LPK2_LOCMI	P11499 locusta mig
9	18	27.7	10	1	TPIS_NICPL	P29.38 nicotiana p
10	17	26.2	7	1	CIA_ENTFA	P16... enterococcu
11	17	26.2	13	1	TEMA_RANTE	P55617 rana temper
12	16	24.6	10	1	ALP9_CARMA	P81822 carcinus ma
13	16	24.6	12	1	DCML_PSEFA	P22... pseudomonas
14	15	23.1	8	1	LPK_LEUMA	P15447 leucophaea
15	15	23.1	9	1	CCAP_CARMA	P38856 carcinus ma
16	15	23.1	9	1	DI_NEPNO	P24816 nephrops no
17	15	23.1	10	1	SLAP_BACTG	P49529 bacillus th
18	15	23.1	11	1	TKNA_SCYCA	P41333 scyliorhinu
19	15	23.1	13	1	UN02_PINPS	P49.. pinus pinas
20	14	21.5	7	1	FAR4_PANRE	P41875 panarellus
21	14	21.5	7	1	TY51_LITRU	P49.. litoria rub
22	14	21.5	8	1	FAR1_PANRE	P41872 panarellus
23	14	21.5	9	1	FAR2_PANRE	P41873 panarellus
24	14	21.5	10	1	HTF1_ROMMI	P10110 romalea mic
25	14	21.5	10	1	HTF2_CARMO	P35.. carausius m
26	14	21.5	10	1	KLA2_MOUSE	P79027 mus musculu
27	14	21.5	10	1	TRT4_LOCMI	P49.. locusta mig
28	14	21.5	10	1	TRT8_LEUMA	P81.. leucophaea
29	14	21.5	10	1	TRT9_LEUMA	P81.. leucophaea
30	14	21.5	10	1	UPA9_HUMAN	P30.. homo sapien
31	14	21.5	11	1	CORZ_PERAM	P11.. periplaneta
32	14	21.5	12	1	LOSK_LOCMI	P47.. locusta mig
33	14	21.5	13	1	CRBL_VESMA	P17... vespa manda
34	14	21.5	13	1	CRBL_VESSI	P17... vespa orien
35	14	21.5	13	1	FAR6_ASCSU	P43173 ascaris su.
36	14	21.5	13	1	TEMP_RANTE	P56.. rana temper
37	13	20.0	8	1	AL11_CARMA	P81.. carcinus ma
38	13	20.0	8	1	AL15_CARMA	P81818 carcinus ma
39	13	20.0	8	1	AL17_CARMA	P81.. carcinus ma
40	13	20.0	8	1	AL18_CARMA	P81821 carcinus ma
41	13	20.0	8	1	AL19_CYBSO	P81.. cybus per s
42	13	20.0	8	1	AL14_CALVI	P1840. calliphora
43	13	20.0	8	1	AL16_CYBSP	P81.. cybus per s
44	13	20.0	8	1	FAR7_ASCSU	P43171 ascaris su.
45	13	20.0	8	1	LCK1_LEUMA	P21.. leucophaea

ALIGNMENTS

RESULT 1
V25K_WSSV
ID V25K_WSSV STANDARD; PRT; 12 AA.
AC P82004;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last Sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE 25 kDa structural polyprotein (Fragment).
OS White spot syndrome virus (WSSV).
OC Viruses; Unassigned viruses.
OX NCBI_TaxID=92652;
RN [1]
RP SEQUENCE.
RC STRAIN=South Carolina;
RX MEDLINE=20214217; PubMed=10795252;
RA Wang Q., Poulos B.T., Lightner D.V.;
RT "Protein analysis of geographic isolates of shrimp white spot syndrome
RT virus.";
RL Arch. Virol. 145:263-274(2000).
CC -!- FUNCTION: STRUCTURAL COMPONENT OF THE VIRION.
FT NON_TER 12 12
SQ SEQUENCE 12 AA; 1263 MW; C54C9AD0E7B71A9 CRC64;

Query Match 68.8%; Score 29; DB 1; Length 12;
Best Local Similarity 80.0%; Pred. No. 4.8e-02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 6 SFILA 10
|||||
Db 4 SFILS 8

RESULT 2
CXAA_CONST
ID CXAA_CONST STANDARD; PRT; 13 AA.
AC P28878;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last Sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Alpha conotoxin SIA (SIA).
OS Conus striatus (Striated cone).
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Caenogastropoda;
OC Neogastropoda; Conoidea; Conidae; Conus.
OX NCBI_TaxID=6493;
RN [1]
RP SEQUENCE.
RC TISSUE=Venom;
RX MEDLINE=91369955; PubMed=1892838;
RA Myers R.A., Zafaralla G.C., Gray W.R., Abbot J., Yoshikami D.,
RA Olivera B.M.;
RT "Alpha-conotoxins, small peptide probes of nicotinic acetylcholine
RT receptors.";
RL Biochemistry 30:9370-9377(1991).
CC -!- FUNCTION: ALPHA CONOTOXINS ACT ON POSTSYNAPTIC MEMBRANES. THEY
CC BIND TO THE NICOTINIC ACETYLCHOLINE RECEPTOR (NACHR) AND INHIB

```
CC       INHIBIT THEM.
CC     -!- SUBCELLULAR LOCATION: Secreted.
CC     -!- SIMILARITY: BELONGS TO THE ALPHA 2 SUPERFAMILY OF CONOTOXIN FAMILY.
DR   PIR; A41783; A41783.
DR   PROSITE; PS00271; CONOTOXIN.
KW   Postsynaptic neurotoxin; Acetylcholine receptor inhibitor; Amidation.
KW   Venom.
FT   DISULFID       2       7
FT   DISULFID       3      13
FT   MOD_RES       13      13       AMIDATION.
SQ   SEQUENCE   13 AA;  1462 MW;  DEF193F0E245788C CRC64;


  Query Match             40.8%;  Score 20;  DB 1;  Length 13;
  Best Local Similarity   42.9%;  Pred. No. 5.2e+02;
  Matches    3;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 CPEGRSF 7
           :  ||
Db       7 CAYGKNF 13


RESULT   4
FARP_CALSI
ID   FARP_CALSI     STANDARD;      PRT;     9 AA.
AC   P38495;
DT   01-OCT-1994 (Rel. 30, Created)
DT   01-OCT-1994 (Rel. 30, Last sequence update)
DT   01-NOV-1995 (Rel. 32, Last annotation update)
DE   FMRFamide like neuropeptide.
OS   Callinectes sapidus (Blue crab).
OC   Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC   Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Brachyura;
OC   Eubrachyura; Portunoidea; Portunidae; Callinectes.
OX   NCBI_TaxID=6763;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE; 92270479; PubMed=1617216;
RA   Kellett K.G.;
RT   "The identification and structure activity relations of a
RT   cardioactive FMRFamide-related peptide from the blue crab Callinectes
RT   sapidus.";
RL   Peptides 12:1295-1302(1991).
CC   -!- FUNCTION: CARDIOACTIVE PEPTIDE.
CC   -!- SIMILARITY: BELONGS TO THE FARP (FMRFAMIDE RELATED PEPTIDE)
CC       FAMILY.
KW   Neuropeptide; Amidation.
FT   MOD_RES       9       9       AMIDATION.
SQ   SEQUENCE   9 AA;  1159 MW;  1548E072905A4045B CRC64;


  Query Match             27.7%;  Score 18;  DB 1;  Length 9;
  Best Local Similarity   60.0%;  Pred. No. 1e+05;
  Matches    3;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       4 FGRSF 7
           : |||
Db       2 YNRSF 6
```

RESULT 6
AKHX_LOCMI
ID AKHX_LOCMI STANDARD; PRT; 10 AA.
AC P41488;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE Locustapyrokinin 2 (LOM-PK-2) (PXPRL amide).
OS Locusta migratoria (Migratory locust).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Orthopteroidea; Orthoptera; Caelifera;
OC Acridomorpha; Acridoidea; Acrididae; Locusta.
OX NCBI_TaxID=7004;
RN [1]
RP SEQUENCE.
RC TISSUE=Corpora cardiaca;
RA Stewart K.J.;
RL Submitted (DEC-1998) to the SWISS-PROT data bank.
CC -!- FUNCTION: PROBABLY INVOLVED IN THE REGULATION OF LIPID
CC INTERMEDIARY METABOLISM, BEHAVIOR AND/OR DEVELOPMENT.
CC -!- SIMILARITY: BELONGS TO THE PYROKININ FAMILY.
DR InterPro; IPR000347; AKH.
KW Neuropeptide; Amidation.
FT MOD_RES 1 1 PYROGLUTAMIC ACID.
FT MOD_RES 10 10 AMIDATION.
SQ SEQUENCE 10 AA; 1222 MW; 81BFF67AB415B9D1 CRC64;

Query Match 27.7%; Score 18; DB 1; Length 10;
Best Local Similarity 42.9%; Prod. No. 9.5e+02;
Matches 3; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 2 IPSRSFI 8
|| : :
Db 3 IESEAWS 9

RESULT 7
O_XM_RAT
ID O_XM_RAT STANDARD; PRT; 10 AA.
AC P04451;
DT 01-NOV-1990 (Rel. 12, Created)
DT 01-NOV-1990 (Rel. 12, Last sequence update)
DT 01-MAY-2002 (Rel. 41, Last annotation update)
DE Cytochrome c oxidase polypeptide VIIb, mitochondrial (EC 1.9.3.1)
DE (Fragment).
GN COX7B.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE.
RC STRAIN=WISTAR; TISSUE=Liver;
RX MEDLINE=95424529; PubMed=7601135;
RA Seppet-Saks M., Novak P., Bolanek W., Brandt U., Saks V.;
RT "Cytochrome c oxidase in developing rat heart. Enzymic properties and
RT amino-terminal sequences suggest identity of the fetal heart and the
RT adult liver isoforms.";
RL Eur. J. Biochem. 230:235-241(1995).
CC -!- FUNCTION: THIS PROTEIN IS ONE OF THE NUCLEAR-CODED POLYPEPTIDE
CC CHAINS OF CYTOCHROME C OXIDASE, THE TERMINAL OXIDASE IN
CC MITOCHONDRIAL ELECTRON TRANSPORT.
CC -!- CATALYTIC ACTIVITY: 4 ferrocytochrome c + O(2) = 4 ferricytochrome
CC c + 2 H(2)O.
KW Oxidoreductase; Mitochondrion.
FT NON_TER 10 10
SQ SEQUENCE 10 AA; 1210 MW; CF070EB771A3332F CRC64;

Query Match 27.7%; Score 18; DB 1; Length 10;
Best Local Similarity 100.0%; Prod. No. 9.5e+02;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PIF 3
||| |
Db 7 PIF 9

RESULT 8
LPKJ_LOCMI
ID LPKJ_LOCMI STANDARD; PRT; 10 AA.
AC P41488;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE Locustapyrokinin 2 (LOM-PK-2) (PXPRL amide).
OS Locusta migratoria (Migratory locust).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Orthopteroidea; Orthoptera; Caelifera;
OC Acridomorpha; Acridoidea; Acrididae; Locusta.
OX NCBI_TaxID=7004;
RN [1]
RP SEQUENCE.
RC TISSUE=Brain;
RX MEDLINE=94094539; PubMed=7903906;
RA Schoofs L., Holman G.M., Nachman R., Proost P., Van Damme J.,
RA de Loof A.;
RT "Isolation, identification and synthesis of locustapyrokinin II (Lom-
RT Locusta migratoria, another member of the PXPRL amide peptide
RT family.";
RL Comp. Biochem. Physiol. 106C:105-109(1993).
CC -!- FUNCTION: MEDIATES VISCERAL MUSCLE CONTRACTILE ACTIVITY (MS ASSAY
CC ACTIVITY).
CC -!- SIMILARITY: BELONGS TO THE PYROKININ FAMILY.
DR InterPro; IPR001484; Pyrokinin.
DR PROSITE; PS00539; PYROKININ; 1.
KW Neuropeptide; Amidation; Pyrokinin.
FT MOD_RES 1 1 PYROGLUTAMIC ACID.
FT MOD_RES 10 10 AMIDATION.
SQ SEQUENCE 10 AA; 1145 MW; CFAF42A1A9CIB272 CRC64;

Query Match 27.7%; Score 18; DB 1; Length 10;
Best Local Similarity 100.0%; Prod. No. 9.5e+02;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PIF 3
||| |
Db 4 PIF 6

RESULT 9
TPIS_NICPL
ID TPIS_NICPL STANDARD; PRT; 10 AA.
AC P19118;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 01-FEB-1996 (Rel. 33, Last annotation update)
DE Triosephosphate isomerase, cytosolic (EC 5.3.1.1) (TIM) (Fragment).
OS Nicotiana plumbaginifolia (leadwort leaved tobacco).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicots,lobeus; core eudicots;
OC Asteridae; euasterids 1; Solanales; Solanaceae; Nicotiana.
OX NCBI_TaxID=4092;
RN [1]
RP SEQUENCE.
RA Rans G., de Jaeger M., Dujon I., Van Montagu M., Van Montagu M.;
RT "Alterations in the phenotype of plant cells selected by NH2 terminal
RT amino acid sequence analysis of proteins electroblotted from two
RT dimensional gel separated total extracts.";
RL Proc. Natl. Acad. Sci. U.S.A. 84:4806-4810(1987).
CC -!- CATALYTIC ACTIVITY: D-glyceraldehyde 3-phosphate = dihydroxyacetone
CC phosphate.

```
CC   -!- PATHWAY: PLAYS AN IMPORTANT ROLE IN CENTRAL CARBON METABOLISM.
CC   -!- SUBUNIT: HOMODIMER.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC   -!- MISCELLANEOUS: IN PLANTS, THERE ARE TWO TYPES OF TPI, CYTOSOLIC
CC       AND PLASTIDIC.
CC   -!- SIMILARITY: BELONGS TO THE TRIOSEPHOSPHATE ISOMERASE FAMILY.
DR   EMBL; AJ25617; AJ27617.
DR   InterPro; IPR000652; Triosephosphate isomerase.
DR   Pfam; PF00121; TIM; 1.
DR   PROSITE; PS00171; TIM_1; PARTIAL.
KW   Glycolysis; Gluconeogenesis; Fatty acid biosynthesis;
KW   Cytoplasm; Probable.
FT   NON_TER       1       1
SQ   SEQUENCE   1 AA;  111 MW;  8A8CC7B52C0C6D1 CRC64;


  Query Match             27.7%;  Score 18;  DB 1;  Length 1;
  Best Local Similarity   75.0%;  Pred. No. 0.65;
  Matches    3;  Conservative    1;  Mismatches    0;  Indels    1;  Gaps    0;

Qy       1 SKCI 4
             :
Db       1 SRFI 4


RESULT  2
ALIA_CARMA
ID   ALIA_CARMA     STANDARD;      PRT;     4 AA.
AC   P11822;
DT   01-JUL-1992 (Rel. 12, Created)
DT   01-JUL-1992 (Rel. 12, Last sequence update)
DT   15-DEC-1995 (Rel. 32, Last annotation update)
DE   Allatostatin precursor (Clumping inducing agent) (CIA).
OS   Carcinus maenas (Common shore crab) (Green crab).
OC   Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC   Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Brachyura;
OC   Eubrachyura; Portunoidea; Portunidae; Carcinus.
OX   NCBI_TaxID=6759;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Cerebral ganglion, and thoracic ganglion;
RX   MEDLINE=98121193; PubMed=9461295;
RA   Duve H., Johnsen A.H., Maestro J. L., Scott A.G., Thorpe P.D.,
RA   Thorpe A.;
RT   "Isolation and identification of multiple neuropeptides of the
RT   allatostatin superfamily in the shore crab Carcinus maenas.";
RL   Eur. J. Biochem. 250:727-734(1997).
CC   -!- FUNCTION: MAY ACT AS A NEUROTRANSMITTER OR NEUROHORMONE.
CC   -!- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.
KW   Neuropeptide; Amidation; Multigene family.
FT   MOD_RES       4       4       AMIDATION (POTENTIAL).
SQ   SEQUENCE   4 AA;  411 MW;  9B6B2C15AB569A81 CRC64;


  Query Match             24.8%;  Score 16;  DB 1;  Length 4;
  Best Local Similarity   95.0%;  Pred. No. 2.6e+03;
  Matches    5;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 PIPGRSFIL 9
             :  |||| |
Db       2 PIDMYSPGL 10


RESULT  3
DCML_PSECA
ID   DCML_PSECA     STANDARD;      PRT;    12 AA.
AC   P19919;
```

```
DT   01-FEB-1991 (Rel. 17, Created)
DT   01-FEB-1991 (Rel. 17, Last sequence update)
DT   01-JUN-1994 (Rel. 29, Last annotation update)
DE   Carbon monoxide oxydehydrogenase (cytochrome B-561) large chain (EC ...)
DE   (Fragment).
OS   Pseudomonas carboxydovorans.
OC   Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC   Bradyrhizobium group; Oligotropha.
OX   NCBI_TaxID=...;
RN   [1]
RP   SEQUENCE.
RC   STRAIN=...;
RX   MEDLINE=...; PubMed=...;
RA   Kraft M., Bredendieck I., Berwin S., Meyer O.;
RT   "Hor loxy and distribution of CO dehydrogenase structural genes in
RT   carboxydotrophic bacteria.";
RL   Arch. Microbiol. 15..:335-341(1989).
CC   -!- CATALYTIC ACTIVITY: CO + H(2)O + ferricytochrome b-561 = CO(2) + 2
CC       H(+) + ferricytochrome b-561.
CC   -!- COFACTOR: Molybdenum.
CC   -!- SUBUNIT: CONSISTS OF THREE POLYPEPTIDE CHAINS: LARGE, MEDIUM AND
CC       SMALL.
DR   PIR; PL0138.
KW   Oxidoreductase; Molybdenum.
FT   VARIANT      12     12       GE -> EK.
FT   N-N IS       12     12
SQ   SEQUENCE   12 AA;  1289 MW;  9617A6768B5B021 CRC64;


  Query Match             24.0%;  Score 15;  DB 1;  Length ...
  Best Local Similarity   75.0%;  Pred. No. 2.8e+05;
  Matches    3;  Conservative    0;  Mismatches    1;  Indels    ;  Gaps    0;

Qy       1 PIPG 4
           |
Db       9 PIAI 12


RESULT  14
IPK_LEUMA
ID   IPK_LEUMA      STANDARD;      PRT;     8 AA.
AC   P14944;
DT   01-JAN-1990 (Rel. 13, Created)
DT   01-FEB-1994 (Rel. 28, Last sequence update)
DT   01-FEB-1994 (Rel. 28, Last annotation update)
DE   Leukopyrokinin (IPK) (LEM-PK).
OS   Leucophaea maderae (Madeira cockroach).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Orthopteroidea; Dictyoptera; Blattaria;
OC   Blaberoidea; Blaberidae; Leucophaea.
OX   NCBI_TaxID=6988;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=...; PubMed=...;
RA   Nachman R.J., Holman G.M., Cook B.J.;
RT   "Active transients and analogs of the insect neuropeptide
RT   leucopyrokinin: structure function studies.",
RL   Biochem. Biophys. Res. Commun. 137:938-942(1986).
RN   [2]
RP   SEQUENCE, AND SYNTHESIS.
RC   TISSUE=Head;
RX   MEDLINE=...; PubMed=...;
RA   Holman G.M., Cook B.J., Nachman R.J.;
RT   "Primary structure and synthesis of a biocket myotropic
RT   neuropeptide isolated from the cockroach, Leucophaea maderae.";
RL   Comp. Biochem. Physiol. 85C:219-224(1986).
CC   -!- FUNCTION: NEUROHORMONE VISCERAL MUSCLE MYOTROPIC ACTIVITY
CC       (MYOSTIMULATORY ACTIVITY).
CC   -!- MISCELLANEOUS: AN ANALOG WITHOUT THE N-TERMINAL PCA RESIDUE WAS
CC       SYNTHESIZED AND FOUND TO EXHIBIT GREATER ACTIVITY (344%) THAN THE
CC       PARENT NEUROPEPTIDE. THE PORTION OF THE SEQUENCE OF IPK M-22
CC       CRITICAL FOR THE MYOTROPIC PROPERTIES IS LIMITED T THE
CC       PENTAPEPTIDE FRAGMENT PIPKI.
CC   -!- SIMILARITY: BELONGS TO THE PYROKININ FAMILY.
DR   PIR; A23967; A23967.
DR   InterPro; IPR001484; Pyrokinin.
DR   PROSITE; PS00539; PYROKININ; 1.
KW   Neuropeptide; Amidation; Pyrokinin.
FT   MOD_RES       1      1       PYROGLU-NE CARBOXYLIC ACI .
FT   MOD_RES       8      8       AMIDATION.
SQ   SEQUENCE   8 AA;  949 MW;  92441771A6D6A1B CRC64;


  Query Match             24.0%;  Score 15;  DB 1;  Length 8
  Best Local Similarity   100.0%;  Pred. No. ...;
  Matches    3;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       6 SPI 8
           |||
Db       3 SPI 5


RESULT  15
CCAP_CARMA
ID   CCAP_CARMA      STANDARD;      PRT;     9 AA.
AC   P38556;
DT   01-OCT-1994 (Rel. 30, Created)
DT   01-OCT-1994 (Rel. 30, Last sequence update)
DT   15-DEC-1998 (Rel. 37, Last annotation update)
DE   Cardioactive peptide (CCAP).
OS   Carcinus maenas (Common shore crab) (Green crab).
OS   Manduca sexta (Tobacco hawkmoth) (Tobacco hornworm).
OS   Tenebrio molitor (Yellow mealworm), and
OS   Spodoptera eridania (Southern armyworm).
OC   Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC   Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Brachyura;
OC   Eubrachyura; Portunoidea; Portunidae; Carcinus.
OX   NCBI_TaxID=6759, 7130, 7067, 42547;
RN   [1]
RP   SEQUENCE.
RC   SPECIES C.maenas; TISSUE=Pericardial organs;
RA   Stangier J., Hilbich C., Beyreuther K., Keller R.;
RT   "Unusual cardioactive peptide (CCAP) from pericardial organs of the
RT   shore crab Carcinus maenas.";
RL   Proc. Natl. Acad. Sci. U.S.A. 84:575-579(1987).
RN   [2]
RP   SEQUENCE.
RC   SPECIES M.sexta;
RX   MEDLINE=92050243; PubMed=1426284;
RA   Cheung C.C., Loi P.K., Sylwester A.W., Lee T.D., Tublitz N.J.;
RT   "Primary structure of a cardioactive neuropeptide from the tobacco
RT   hawkmoth, Manduca sexta.";
RL   FEBS Lett. 313:165-168(1992).
RN   [3]
RP   SEQUENCE.
RC   SPECIES T.molitor, and S.eridania; TISSUE=Head;
RX   MEDLINE=94176032; PubMed=8112...;
RA   Furuya K., Liao S., Reynolds S.E., Ota K.B., Hackett M.
RA   Schooley D.A.;
RT   "Isolation and identification of a cardioactive peptide from Tenebrio
RT   molitor and Spodoptera eridania.";
RL   Biol. Chem. Hoppe-Seyler 374:1295-1074(1993).
CC   -!- FUNCTION: THE EFFECT OF CCAP IS BOTH IN- AND MODULATORY.
CC   -!- TISSUE SPECIFICITY: STORED IN PERICARDIAL ORGANS AND RELEASED
CC       INTO THE HEMOLYMPH.
DR   PIR; A26363; A26363.
DR   PIR; S27233; S27233.
KW   Neuropeptide; Amidation.
FT   DISULFID      3      9
FT   MOD_RES       9      9       AMIDATION.
SQ   SEQUENCE   9 AA;  959 MW;  05A0c1A7C1433ec CRC64;


  Query Match             24.0%;  Score 15;  DB 1;  Length 9
  Best Local Similarity   90.0%;  Pred. No. ...e-05;
  Matches    3;  Conservative    1;  Mismatches    ...;  Indels    ;  Gaps    ;
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

SW protein: protein search, using sw model

Run on: August 26, 2002, 06:12:53 ; Search time 81.27 Seconds
(without alignments)
27.449 Million cell updates/sec

Title: US-09-744-282-5
Perfect score: 65
Sequence: 1 PTPGRSFTLASSE 13

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 2.0

Minimum DB seq length: 2
Maximum DB seq length: 13

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: SPTREMBL_19:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	24	36.9	10	10	Q9S905	Q9S905 glycine max
2	24	36.9	13	12	Q64813	Q64813 autographa
3	20	30.8	9	15	Q85710	Q85710 bos taurus
4	20	30.8	12	7	Q31006	Q31006 bos taurus
5	20	30.8	13	10	Q9S922	Q9S922 brassica
6	19	29.2	8	2	Q9R5L7	Q9R5L7 crocodylus
7	19	29.2	13	4	Q15537	Q15537 homo sapien
8	18	27.7	11	2	Q94M17	Q94M17 escherichia
9	18	27.7	11	5	P82698	P82698 tenopharea
10	18	27.7	12	4	Q9P116	Q9P116 homo sapien
11	18	27.7	12	6	Q9TR01	Q9TR01 bos taurus
12	18	27.7	13	2	Q9KBI4	Q9KBI4 caulobacter
13	17	26.2	8	2	Q9S6D5	Q9S6D5 escherichia
14	17	26.2	8	3	P82858	P82858 puccinia re
15	17	26.2	9	10	Q9AXH8	Q9AXH8 mesembryant
16	17	26.2	10	13	Q9PR09	Q9PR09 sparus aura
17	17	26.2	11	2	Q9EUZ8	Q9EUZ8 escherichia
18	17	26.2	12	2	Q50442	Q50442 mycoplasma
19	17	26.2	13	4	O25085	O25085 homo sapien
20	17	26.2	13	4	Q9HAQ9	Q9HAQ9 homo sapien
21	17	26.2	13	5	Q90513	Q90513 trypanosoma
22	16	24.6	8	2	P72221	P72221 pseudomonas
23	16	24.6	9	2	Q96560	Q96560 salmonella
24	16	24.6	9	4	Q15009	Q15009 homo sapien
25	16	24.6	9	6	Q9XSL0	Q9XSL0 capra hircu
26	16	24.6	9	11	Q9QVB9	Q9QVB9 mus sp. sp
27	16	24.6	10	6	Q99M70	Q99M70 rattus norv
28	16	24.6	13	6	Q91KW6	Q91KW6 bos taurus
29	15.5	23.8	11	6	Q9GL48	Q9GL48 sus scrofa
30	15	23.1	7	2	P70804	P70804 acetobacter
31	15	23.1	8	8	P92422	P92422 psathyrella
32	15	23.1	8	8	P92473	P92473 lactarius t
33	15	23.1	8	8	P93985	P93985 aeolidia p
34	15	23.1	8	8	P92404	P92404 lepbeptyma
35	15	23.1	8	8	P92429	P92429 psendocon
36	15	23.1	8	8	P93973	P93973 cromopyrm
37	15	23.1	8	8	P93970	P93970 cromopyrm
38	15	23.1	8	8	P92388	P92388 betaretina
39	15	23.1	8	8	P92428	P92428 petalonema
40	15	23.1	8	8	P92391	P92391 heteranthel
41	15	23.1	8	8	P92227	P92227 erithopsis
42	15	23.1	8	8	P93968	P93968 psathyrella
43	15	23.1	8	8	P93961	P93961 psathyrella
44	15	23.1	8	8	P92215	P92215 amboycyprum
45	15	23.1	8	8	P93981	P93981 erithedrum

ALIGNMENTS

RESULT 1
Q9S905
ID Q9S905 PRELIMINARY; PRT; 10 AA.
AC Q9S905;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE S3 PEPTIDE.
OS Glycine max (Soybean).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
OX NCBI_TaxID=3847;
RN [1]
RP SEQUENCE.
RX MEDLINE=92232221; PubMed=1568167;
RA Hirano H., Kasawa H., Okubo K.;
RL Phytochemistry 31:731-735(1992).
SQ SEQUENCE 10 AA; 975 MW; 47744271AA74CB7 CRC64;

Query Match 36.9%; Score 24; DB 10; Length 10
Best Local Similarity 100.0%; Pred. No. 4.1e+0
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0

Qy 1 PTPG 4
||||
Db 3 PTPG 6

RESULT 2
Q64813
ID Q64813 PRELIMINARY; PRT; 13 AA.
AC Q64813;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE POLYHEDRIN (FRAGMENT).
OS Autographa californica nuclear polyhedrosis virus (AcMNPV).

```
  ;   Viruses; dsDNA viruses, no RNA stage; Baculoviridae; Eu
  ;   Nucleopolyhedrovirus.
RX   NCBI_TaxID=46015;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=?;
RX   MEDLINE=94226764; PubMed=8112993;
RA   Li Y.-S., Howard S.C.;
RT   "Analysis of the polyhedrin gene promoter of the Autographa
RT   californica nuclear polyhedrosis virus.";
RL   Nucleic Acids Res. 15:10233-10248(1987).
DR   EMBL; X66477; CAA46946.1; -.
FT   NON_TER       1       1
SQ   SEQUENCE   12 AA;  1557 MW;  1A25216D0EBE741B CRC64;


  Query Match             30.8%;  Score 24;  DB 12;  Length 12;
  Best Local Similarity   66.7%;  Pred. No. 5.5e+02;
  Matches    4;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 FIFSRS 6
           |: |:|
Db       4 FIISRI 10

RESULT   4
Q9C7J0
ID   Q9C7J0      PRELIMINARY;      PRT;     9 AA.
AC   Q9C7J0;
DT   01-NOV-1999 (TrEMBLrel. 01, Created)
DT   01-NOV-1999 (TrEMBLrel. 01, Last sequence update)
DT   01-NOV-2000 (TrEMBLrel. 08, Last annotation update)
DE   10 KDA PROTEIN (FRAGMENT).
OS   Bacteriophage T4.
OC   Viruses; ds-DNA viruses, Retroviridae; Alpharetrovirus;
OX   NCBI_TaxID=12896;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=84211040; PubMed=6419754;
RA   Sekiya T., Hiratsuka H.;
RT   "DNA sequence of the Bryan high titer strain of Rous sarcoma virus:
RT   deletion and possible genetical relationship with
RT   other strains.";
RL   J. Virol. 49:549-556(1984).
DR   EMBL; K00441; AAA42567.1;
FT   NON_TER       1       1
SQ   SEQUENCE   9 AA;  949 MW;  91AA144E6D7C1AA CRC64;


  Query Match             30.8%;  Score 20;  DB 5;  Length 9;
  Best Local Similarity   66.7%;  Pred. No. 5.5e+02;
  Matches    4;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       8 IIASSE 13
           |:|:|:
Db       3 IISSIE 8


RESULT   5
Q9C7K1
ID   Q9C7K1      PRELIMINARY;      PRT;    12 AA.
AC   Q9C7K1;
DT   01-NOV-1999 (TrEMBLrel. 01, Created)
DT   01-NOV-1999 (TrEMBLrel. 01, Last sequence update)
DT   01-DEC-2000 (TrEMBLrel. 10, Last annotation update)
DE   90A PROTEIN (FRAGMENT).
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini;
OC   Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=93052564; PubMed=1428011;
RA   Ellis S.A., Braem K.A., Morrison W.I.;
RT   "Transmembrane and cytoplasmic domain sequences demonstrate at least
RT   two expressed bovine MHC class I loci.";
RL   Immunogenetics 37:49-56(1992).
DR   EMBL; S42738; AAB23972.1; -.
FT   NON_TER       1       1
SQ   SEQUENCE   12 AA;  1406 MW;  6D9E2FB05AB85044 CRC64;


  Query Match             30.8%;  Score 20;  DB 7;  Length 12;
  Best Local Similarity   50.0%;  Pred. No. 5e+03;
  Matches    4;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       4 GRSFILAS 11
           || :| |:
Db       4 GRIYIQAA 11


RESULT   5
Q9S922
ID   Q9S922      PRELIMINARY;      PRT;    13 AA.
AC   Q9S922;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   S3 SGP S LOCUS SPECIFIC GLYCOPROTEIN.
OS   Brassica oleracea (Cauliflower).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Brassicales; Brassicaceae; Brassica.
OX   NCBI_TaxID=3712;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=92090397; PubMed=1752245;
RA   Gaude T., Demoroy L., Dumas C.;
RT   "Use of a fast protein electrophoretic purification procedure for N
RT   terminal sequence analysis to identify S locus related proteins in
RT   stigmas of Brassica oleracea.";
RL   Electrophoresis 12:646-653(1991).
SQ   SEQUENCE   13 AA;  1351 MW;  9023EDFE99960185 CRC64;


  Query Match             30.8%;  Score 20;  DB 10;  Length 13;
  Best Local Similarity   66.7%;  Pred. No. 5.2e+03;
  Matches    4;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       8 IIASSE 13
           |:|:|:
Db       3 IISSIE 8


RESULT   6
Q9K5L7
ID   Q9K5L7      PRELIMINARY;      PRT;     8 AA.
AC   Q9K5L7;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE   1,4 BETA D GLUCAN GLUCANOHYDROLASE (EC 3.2.1.4) (FRAGMENT).
OS   Clostridium thermocellum.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
OC   Clostridium.
OX   NCBI_TaxID=1515;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=92231850; PubMed=1567379;
RA   Romaniec M.P., Faath U., Kobayashi T., Baskisson N.S., Barker P.J.,
RA   Demain A.L.;
RT   "Purification and characterization of a new endoglucanase from
RT   Clostridium thermocellum.";
RL   Biochem. J. 283:69-73(1992).
```

SQ SEQUENCE 8 AA; 823 MW; C2C1AB1B9B9B1B775 CRC64;

Query Match 29.2%; Score 19; DB 2; Length 8;
Best Local Similarity 50.0%; Pred. No. 5.6e+05;
Matches 3; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 FPFPSS 6
Db 2 FIPAPT 7

RESULT 7
Q16687
ID Q16687 PRELIMINARY; PRT; 10 AA.
AC Q16687;
DT 01-NOV-1999 (IrEMBLrel. 01, Created)
DT 01-NOV-1999 (IrEMBLrel. 01, Last sequence update)
DT 01-JAN-1999 (IrEMBLrel. 09, Last annotation update)
DE DNA FROM CHROMOSOME 9q, CONTAINING TRANSLOCATION BREAKPOINT.
DE (FRAGMENT).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95167754; PubMed=7892812;
RA Fujino M., Takayama N., Kato S., Matsuno K., Takano I., Nakazawa S.;
RT "Rearrangement of immunoglobulin kappa gene in a variant type Burkitt
RT lymphoma cell line. Kappa101, carrying a translocation between
RT chromosomes 2 and 8.";
RL Jpn. J. Cancer Res. 77:24-27(1986).
RN [2]
RP SEQUENCE FROM N.A.
RA Kato S.;
RL Submitted (SEP-1990) to the EMBL/GenBank/DDBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Takano I.;
RL Submitted (SEP-1990) to the EMBL/GenBank/DDBJ databases.
DR EMBL; D00160; BAA14191.1; -.
FT NON_TER 1 1
SQ SEQUENCE 10 AA; 1535 MW; 2BC1FE941F9151AB CRC64;

Query Match 29.2%; Score 19; DB 4; Length 10;
Best Local Similarity 50.0%; Pred. No. 5e+05;
Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 2 FPSFTL 9
Db 2 FPGSTRL 9

RESULT 8
Q93MI7
ID Q93MI7 PRELIMINARY; PRT; 11 AA.
AC Q93MI7;
DT 01-DEC-2001 (IrEMBLrel. 19, Created)
DT 01-DEC-2001 (IrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (IrEMBLrel. 19, Last annotation update)
DE ADHESIN (FRAGMENT).
GN CHA.
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CFT073;
RA Stoll A.J.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AE017502; AAK94916.1; -.
FT NON_TER 11 11
SQ SEQUENCE 11 AA; 1203 MW; 8E26C7E9B1D7C1BC CRC64;

Query Match 27.7%; Score 18; DB 2; Length 11;
Best Local Similarity 71.4%; Pred. No. 2.5e+05;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 5 RSFTLAS 11
Db 2 RTITLAS 8

RESULT 9
P82698
ID P82698 PRELIMINARY; PRT; 11 AA.
AC P82698;
DT 01-MAR-2001 (IrEMBLrel. 16, Created)
DT 01-MAR-2001 (IrEMBLrel. 16, Last sequence update)
DT 01-JUN-2001 (IrEMBLrel. 17, Last annotation update)
DE PERIVISCEROKININ 1 (LEM-PVK-1).
OS Leucophaea maderae (Madeira cockroach).
OS Nauphoeta cinerea (Cinereous cockroach) (Gray cockroach).
OS Blaberus craniifer,
OS Blaptica dubia, and
OS Gromphadorina portentosa (Cockroach).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Orthopteroidea; Dictyoptera; Blattaria;
OC Blaberoidea; Blaberidae; Leucophaea.
OX NCBI_TaxID=6988, 6990, 6982, 132935, 36953;
RN [1]
RP SEQUENCE, FUNCTION, AND MASS SPECTROSCOPY.
RC TISSUE=ABDOMINAL PERISYMPATHETIC ORGANS;
RX MEDLINE=20307624; PubMed=10844006;
RA Predel R., Kellner R., Baggerman G., Steinmetzer ., Schoofs L.;
RT "Upregulation of novel perisinsec kinins in single neurohemal
RT release sites in insects: MS/MS fragmentation complemented by Edman
RT degradation.";
RL Eur. J. Biochem. 267:3869-3873(2000).
CC -!- FUNCTION: MYOTROPIC NEUROPEPTIDE WITH MYOTROPIC ACTIVITY
CC (MYOTROPIC ACTIVITY).
CC -!- MASS SPECTROMETRY: MW 1090.6, METHOD MALDI.
KW Neuropeptide; Amidation.
FT MOD_RES 11 11 AMIDATION.
SQ SEQUENCE 11 AA; 1091 MW; 2C2D80E227C05220 CRC64;

Query Match 27.7%; Score 18; DB 5; Length 11;
Best Local Similarity 75.0%; Pred. No. 6.2e+05;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 3 FGKS 6
Db 8 FGRT 11

RESULT 10
Q9PI16
ID Q9PI16 PRELIMINARY; PRT; 11 AA.
AC Q9PI16;
DT 01-OCT-2000 (IrEMBLrel. 15, Created)
DT 01-OCT-2000 (IrEMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (IrEMBLrel. 15, Last annotation update)
DE APOLIPOPROTEIN E RECEPTOR 2 (FRAGMENT).
OS Homo Sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Korschineck I., Gharehbaghi Schnell E., Lang I., Binder B.R.;
RT "Expression of Apolipoprotein E Receptor 2 in Atherosclerotic

```
RL    Submitted (FEB-1999) to the EMBL/GenBank/DDBJ databases.
DR    EMBL; AF124177; AAF06440.1; -.
KW    Exception; Lipoprotein.
FT    NON_TER       1       1
FT    NON_TER       12      12
SQ    SEQUENCE   12 AA;  1467 MW;  4E0935410404187 CRC64;


  Query Match             27.7%;  Score 18;  DB 4;  Length 12;
  Best Local Similarity   100.0%;  Pred. No. 1.6e+03;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 PIPIS 5
           |||||
Db       7 PIPIS 11


RESULT  13
Q9S6D5
ID   Q9S6D5       PRELIMINARY;      PRT;     8 AA.
AC   Q9S6D5;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   PUTATIVE IS30 TRANSPOSASE (FRAGMENT).
OS   Escherichia coli.
OC   Bacteria; Proteobacteria; Gamma subdivision; Enterobacteriaceae;
OC   Escherichia.
OX   NCBI_TaxID=562;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN A295B;
RX   MEDLINE=99194747; PubMed=10094716;
RA   Kahn A., Drummelsmith J., Whitfield C.;
RT   "Conserved organization in the cps gene clusters for expression of
RT   Escherichia coli group 1 K antigens: relationship to the colanic acid
RT   biosynthesis locus and the cps genes from Klebsiella pneumoniae.";
RL   J. Bacteriol. 181:2307-2313(1999).
DR   EMBL; AF118251; AAD30008.1; -.
FT   NON_TER       8       8
SQ   SEQUENCE   8 AA;  1011 MW;  F2110C1A9DIB414D6 CRC64;


  Query Match             26.2%;  Score 17;  DB 2;  Length 8;
  Best Local Similarity   75.0%;  Pred. No. 5.6e+05;
  Matches    6;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       5 RSFI 8
           |:||
Db       3 RIFI 6


RESULT  14
P82858
ID   P82858       PRELIMINARY;      PRT;     8 AA.
AC   P82858;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   Uricase (Fragment) (Urate oxidase) (FRAGMENT).
OS   Puccinia recondita f. sp. tritici.
OC   Eukaryota; Fungi; Basidiomycota; Urediniomycetes; Urediniomycetidae;
OC   Uredinales; Pucciniaceae; Puccinia.
OX   NCBI_TaxID=142079;
RN   [1]
RP   SEQUENCE, FUNCTION, AND DEVELOPMENTAL STAGE.
RC   TISSUE=SPORE;
RA   Aguilar M., Montalbini P., Pineda M.;
RL   Submitted (NOV-2000) to the SWISS-PROT data bank.
CC   -!- FUNCTION: INVOLVED IN HOST-PARASITE RELATIONSHIP BETWEEN PLANTS
CC       AND FUNGI.
CC   -!- CATALYTIC ACTIVITY: URATE + O(2) + H(2)O = 5-HYDROXYISOURATE +
CC       H(2)O(2) (5-HYDROXYISOURATE DECOMPOSES TO FORM ALLANTOIN).
CC   -!- SUBCELLULAR LOCATION: PEROXISOMAL (BY SIMILARITY).
CC   -!- DEVELOPMENTAL STAGE: GERMINATION.
CC   -!- SIMILARITY: BELONGS TO THE URICASE FAMILY.
DR   InterPro; IPR002042; Uricase.
DR   PROSITE; PS00366; URICASE; PARTIAL.
KW   Oxidoreductase; Purine metabolism; Peroxisome.
FT   NON_TER       8       8
SQ   SEQUENCE   8 AA;  777 MW;  98C1ADD735B9D76D CRC64;
```

Query Match 26.2%; Score 17; DB 3; Length 9;
Best Local Similarity 60.0%; Pred. No. 5.6e+05;
Matches 3; Conservative 3; Mismatches 0; Indels 2; Gaps 0;

RESULT 15
Q9AXH8
ID Q9AXH8 PRELIMINARY; PRT; 9 AA.
AC Q9AXH8;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE HIP A_UAi-kIN (FRAGMENT).
GN MIPK.
OS Mesembryanthemum crystallinum (Common ice plant).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Caryophyllidae; Caryophyllales; Aizoaceae; Mesembryanthemum.
OX NCBI_TaxID=3544;
RN [1]
RP SEQUENCE FROM N.A.
RA Bohnert H.J., Quigley F.L., Bohnert H.J.;
RT "The promoter of mipK gene which encodes for a HIP aquaporin from
RT Mesembryanthemum crystallinum.";
RL Submitted (SEP-2000) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AF326727; AAG40597.1; -.
KW Fragment.
FT NON_TER 9 9
SQ SEQUENCE 9 AA; 969 MW; C4EB0B7787D72871 CRC64;

Query Match 26.2%; Score 17; DB 10; Length 9;
Best Local Similarity 60.0%; Pred. No. 5.6e+05;
Matches 3; Conservative 3; Mismatches 0; Indels 2; Gaps 0;

Search completed: August 26, 2002, 06:23:01
Job time : 600 secs

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

M protein, protein search, using sw model

Run on: August 26, 2002, 06:11:33 ; Search time 97.53 Seconds
(without alignments)
17.895 Million cell updates/sec

Title: US-09-744-282-5
Perfect score: 65
Sequence: 1 PTFGRSPTLASSE 13

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0
Maximum DB seq length: 13

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_032802:*

1:	/SIDS2/gcgdata/geneseq/geneseq-pep/	3	25	*
2:	/SIDS2/gcgdata/geneseq/geneseq-pep/	3	25	*
3:	/SIDS2/gcgdata/geneseq/geneseq-pep/	3	25	*
4:	/SIDS2/gcgdata/geneseq/geneseq-pep/	3	25	*
5:	/SIDS2/gcgdata/geneseq/geneseq-pep/	3	25	*
6:	/SIDS2/gcgdata/geneseq/geneseq-pep/	3	25	*
7:	/SIDS2/gcgdata/geneseq/geneseq-pep/	3	25	*
8:	/SIDS2/gcgdata/geneseq/geneseq-pep/	3	25	*
9:	/SIDS2/gcgdata/geneseq/geneseq-pep/	3	25	*
10:	/SIDS2/gcgdata/geneseq/geneseq-pep/	3	25	*
11:	/SIDS2/gcgdata/geneseq/geneseq-pep/	3	25	*
12:	/SIDS2/gcgdata/geneseq/geneseq-pep/	3	25	*
13:	/SIDS2/gcgdata/geneseq/geneseq-pep/	3	25	*
14:	/SIDS2/gcgdata/geneseq/geneseq-pep/	3	25	*
15:	/SIDS2/gcgdata/geneseq/geneseq-pep/	3	25	*
16:	/SIDS2/gcgdata/geneseq/geneseq-pep/	3	25	*
17:	/SIDS2/gcgdata/geneseq/geneseq-pep/	3	25	*
18:	/SIDS2/gcgdata/geneseq/geneseq-pep/	3	25	*
19:	/SIDS2/gcgdata/geneseq/geneseq-pep/	3	25	*
20:	/SIDS2/gcgdata/geneseq/geneseq-pep/	3	25	*
21:	/SIDS2/gcgdata/geneseq/geneseq-pep/	3	25	*
22:	/SIDS2/gcgdata/geneseq/geneseq-pep/	3	25	*

first, No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	65	100.0	13	17	AAR95075	Peptide SEQ ID
2	65	100.0	13	21	AAY69230	Peptide fragment o
3	14	75.4	13	21	AAB03439	Spot# derived cell
4	44	67.7	9	17	AAR95071	Human cartilage gl
5	44	67.7	9	21	AAY69226	Peptide fragment o
6	41	63.1	13	17	AAR95076	Peptide SEQ ID
7	41	63.1	13	19	AAW61597	Human cartilage gp
8	41	63.1	13	21	AAY69241	Peptide fragment o
9	41	63.1	13	21	AAY77401	Human cartilage gl
10	41	63.1	13	22	AAU03045	Human cartilage gp
11	41	63.1	13	22	AAU03048	Human cartilage gp
12	41	63.1	13	22	AAU03049	Human cartilage gp
13	41	63.1	13	22	AAU03050	Human cartilage gp
14	41	63.1	13	22	AAU03063	Human cartilage gp
15	41	63.1	13	22	AAU03065	Human cartilage gp
16	41	63.1	13	22	AAU03066	Human cartilage gp
17	41	63.1	13	22	AAU03067	Human cartilage gp
18	37	56.9	13	22	AAU03054	Human cartilage gp
19	37	56.9	13	22	AAU03055	Human cartilage gp
20	37	56.9	13	22	AAU03073	Human cartilage gp
21	36	55.4	13	22	AAU03046	Human cartilage gp
22	36	55.4	13	22	AAU03047	Human cartilage gp
23	36	55.4	13	22	AAU03051	Human cartilage gp
24	36	55.4	13	22	AAU03052	Human cartilage gp
25	36	55.4	13	22	AAU03057	Human cartilage gp
26	36	55.4	13	22	AAU03058	Human cartilage gp
27	36	55.4	13	22	AAU03062	Human cartilage gp
28	36	55.4	13	22	AAU03064	Human cartilage gp
29	36	55.4	13	22	AAU03072	Human cartilage gp
30	34	52.3	13	22	AAU03053	Human cartilage gp
31	34	52.3	13	22	AAU03059	Human cartilage gp
32	34	52.3	13	22	AAU03060	Human cartilage gp
33	34	52.3	13	22	AAU03071	Human cartilage gp
34	33	50.8	12	21	AAB95067	Human beta-5 subun
35	33	50.8	13	21	AAY77400	Human chondrocyte
36	32	49.2	9	17	AAR95072	Human cartilage gl
37	32	49.2	9	21	AAY69227	Peptide fragment o
38	30	46.2	13	22	AAU03056	Human cartilage gp
39	29	44.6	9	21	AAY77399	Human chondrocyte
40	29	44.6	10	22	ABB55898	Vascular dementia
41	29	44.6	10	22	ABB55939	Vascular dementia
42	29	44.6	11	22	ABB52311	Vascular dementia
43	29	44.6	11	22	ABB52306	Human API 41 trypt
44	29	44.6	13	22	AAU29539	API trypsin digest
45	29	44.6	13	22	AAU28751	API trypsin digest

ALIGNMENTS

RESULT 1
AAR95075
ID AAR95075 standard; peptide; 13 AA.
XX
AC AAR95075;
XX
DT 17-DEC-1996 (first entry)
XX
DE Peptide #1 to be used in production of monoclonal
XX
KW Human cartilage; glycoprotein 39; HC gp 39; autoantigen;
KW rheumatoid arthritis; T-cell tolerance; B-cell tolerance;
KW rheumatoid arthritis; disease model; diagnosis.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Peptide 1..13
FT /note= "HC gp-39 peptide (1)"
XX
PN WO9613517-A1.
XX
PD 09-MAY-1996.
XX
PF 25-OCT-1995; 95WO-EP04201.
XX
PR 07-APR-1995; 95EP-0200886.
PR 27-OCT-1994; 94EP-0203128.
XX
PA (AKZO-) AKZO NOBEL NV.
XX
PI Boots AMH, Verheijden GEM;
XX
DR WPI; 1996-239447/24.

XX
PT Peptide(s) derived from human cartilage gp 39 autoantigen useful
PT to induce specific T cell tolerance, esp. for treatment of
PT rheumatoid arthritis, also useful in producing arthritic animal
PT models.
XX
PS Claim 1; Page 20; 37pp; English.
XX
CC The present peptides contain a peptide derived from the human
CC cartilage glycoprotein 39 (HC gp 39) autoantigen, and can be used to
CC induce specific T cell tolerance to HC gp 39 in a patient with
CC T cell mediated cartilage destruction, esp. rheumatoid arthritis
CC (RA). It may also be used to induce RA in mice, etc. (human disease
CC models), and to diagnose the presence of activated T cells
CC autoreactive with HC gp 39, i.e. to identify patients suitable for
CC tolerising treatment.
CC Therapeutically the peptide is given in a dose of 1-1 100 microg/kg
CC by injection, or 10-50 microg/kg to produce a disease model. The
CC peptide has a specific tolerising effect on autoreactive T cells,
CC but leaves other components of the immune system intact, and
CC therefore should have few side effects.
XX
SQ Sequence 13 AA;

Query Match 100.0%; Score 65; DB 21; Length 13;
Best Local Similarity 100.0%; Pred. No. 7.7e-06;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PTFGRSFTLASSE 13
||||||||||||| |
Db 1 PTFGRSFTLASSE 13

RESULT 3
AAB03439
ID AAB03439 standard; Peptide; 13 AA.
XX
AC AAB03439;
XX
DT 03-JAN-2001 (first entry)
XX
DE Gp38k derived cell migration peptide #5.
XX
KW Gp38k; chemoattractant; cell migration; wound healing; angiogenesis;
KW cancer; vascular trauma; vascular disease; atherosclerosis; restenosis.
XX
OS Unidentified.
XX
PN WO200034469-A1.
XX
PD 15-JUN-2000.
XX
PF 10-DEC-1999; 99WO-US29262.
XX
PR 11-DEC-1998; 98US-0111856.
XX
PA (UYNY) UNIV NEW YORK STATE RES FOUND.
XX
PI Mills A;
XX
DR WPI; 2000-431380/37.
XX
PT Clusterin and gp38K-related peptide capable of arterial cell migration
PT useful for treating atherosclerosis, cancer and stenosis following
PT vascular trauma or disease
XX
PS Claim 1; Page 26; 48pp; English.
XX
CC The present sequence is a peptide derived from gp38k. Gp38k is
CC essential for the migration of vascular smooth muscle cells (VSMC). The
CC gene and protein can, therefore, be used to promote wound healing,
CC angiogenesis and vasculogenesis, in the treatment of stenosis following
CC vascular trauma or disease and to treat atherosclerosis, and antisense
CC sequences can be used to treat cancer, as angiogenesis is vital for
CC tumour survival. This peptide can be used to alter cell migration
CC patterns.
XX
SQ Sequence 13 AA;

Query Match 75.4%; Score 49; DB 21; Length 13;
Best Local Similarity 90.9%; Pred. No. 0.0091;
Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 8 FGRSFILASSE 18
Db 1 tarstflassk 11

RESULT 4
AAR69076
ID AAR69076 standard; peptide; 9 AA.
XX
AC AAR69076;
XX
DT 17-DEC-1996 (first entry)
XX
DE Human cartilage glycoprotein 39 autoantigen derived peptide (1).
XX
KW Human; cartilage; glycoprotein 39; hC gp-39; autoantigen;
KW induction; T-cell; tolerance; mediated; destruction;
KW rheumatoid arthritis; disease model; diagnosis.
XX
OS Homo sapiens.
XX
PN WO9615151-A1.
XX
PD 23-MAY-1996.
XX
PF 25-OCT-1995; 95WO-EP04201.
XX
PR 07-APR-1995; 95EP-0200886.
PR 27-OCT-1994; 94EP-0204128.
XX
PA (ALKU) AKZO NOBEL NV.
XX
PI Boots AMH, Verheijden GFM;
XX
DR WPI; 1996-239447/24.
XX
PT Peptide(s) derived from human cartilage gp-39 autoantigen - used
PT to induce specific T cell tolerance, esp. for treatment of
PT rheumatoid arthritis, also useful in producing arthritic animal
PT models
XX
PS Claim 1; Page 28; 37pp; English.
XX
CC The present peptide, derived from the human cartilage glycoprotein
CC 39 (hC gp-39) autoantigen, can be used to induce specific T cell
CC tolerance to hC gp-39 in a patient with T-cell mediated cartilage
CC destruction, esp. rheumatoid arthritis (RA). It may also be used to
CC induce RA in mice, etc. (human disease models), and to diagnose the
CC presence of activated T cells autoreactive with hC gp-39, i.e. to
CC identify patients suitable for tolerising treatment.
CC Therapeutically the peptide is given in a dose of 0.1-100 micromol/kg
CC by injection, or 10-50 micromol/kg to produce a disease model. The
CC peptide has a specific tolerising effect on autoreactive T-cells,
CC but leaves other components of the immune system intact, and
CC therefore should have few side effects. In an assay to determine
CC the peptide's binding to HLA-DR mols., specifically DR4(w4 (known
CC to be associated with an increased risk of RA development), the
CC peptide gave a Dot blot IC(50) (micromolar) value of 0.04 compared
CC to 0.55 for the control peptide IHA(307-319)F.
XX
SQ Sequence 9 AA;

Query Match 67.7%; Score 44; DB 17; Length 9;
Best Local Similarity 100.0%; Pred. No. 6.4e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 8 FGRSFILAS 11
Db 1 tarstflas 9

RESULT 5
AAY69226
ID AAY69226 standard; peptide; 9 AA.
XX
AC AAY69226;
XX
DT 30-MAY-2000 (first entry)
XX
DE Peptide fragment of human cartilage gp-39 (HC gp-39) protein.
XX
KW Human cartilage gp-39 protein; HC gp-39 protein; T cell activator;
KW inflammatory disease; lymphocyte; antigen; bystander suppression;
KW antigen-specific T cell tolerance; autoimmune disorder;
KW inflammatory disease; rheumatoid arthritis;
KW autoimmune inflammatory disease; autoantigen.
XX
OS Homo sapiens.
XX
PN WO200004917-A2.
XX
PD 03-FEB-2000.
XX
PF 19-JUL-1999; 99WO-EP05331.
XX
PR 23-JUL-1998; 98EP-0202471.
XX
PA (ALKU) AKZO NOBEL NV.
XX
PI Miltenburg AMM, Boots AMH;
XX
DR WPI; 2000-182937/16.
XX
PT Use of cartilage protein or peptides for treating inflammatory diseases
PT especially rheumatoid arthritis -
XX
PS Claim 2; Page 17; 29pp; English.
XX
CC AAY69226-33 represent fragments of the human cartilage gp-39 (HC gp-39)
CC protein. The peptides are T cell activators. The peptides are used for
CC preventing inflammatory diseases, by mediating lymphocytes reactive
CC to antigens other than HC gp-39, present in the same tissue of
CC HC gp-39. By the induction of antigen-specific T-cell tolerance,
CC autoimmune disorders can be treated by bystander suppression. The
CC HC gp-39 peptides are useful for preventing inflammatory diseases
CC especially rheumatoid arthritis. The fragments may also be used for
CC treating autoimmune inflammatory diseases. The HC gp-39 peptides are
CC especially beneficial in autoimmune conditions in which the
CC autoantigens associated with the condition are unknown.
XX
SQ Sequence 9 AA;

Query Match 67.7%; Score 44; DB 21; Length 9;
Best Local Similarity 100.0%; Pred. No. 6.4e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 8 FGRSFILAS 11
Db 1 tarstflas 9

RESULT 6
AAR95076
ID AAR95076 standard; peptide; 18 AA.
XX
AC AAR95076;
XX
DT 17-DEC-1996 (first entry)
XX
DE Peptide contg. hC gp-39 autoantigen derived peptide (21).
XX
KW Human; cartilage; glycoprotein; hC gp-39; autoantigen;
KW induction; T-cell; tolerance; mediated; destruction;
KW rheumatoid arthritis; disease model; diagnosis.

XX
[illegible]
XX
FH Key Location/Qualifiers
FT [illegible] 6..11
FT /note= "HC gp 39 peptide (11)"
XX
PN WO 9535317 A.
XX
PD 28-MAY-1995.
XX
PF 21-DEC-1995; 95WO-EP04201.
XX
PR 21-DEC-1994; 94EP-0203128.
XX
PA (AKZO) AKZO NOBEL NV.
XX
PI Boots AMH, Verheijden GEM;
XX
DR WPI; 1996-23447/24.
XX
PT Peptide(s) derived from human cartilage gp-39 autoantigen - used
PT to induce specific T cell tolerance, esp. for treatment of
PT rheumatoid arthritis, also useful in producing arthritic animal
PT model.
XX
PS Claim 2; Page 28; 37pp; English.
XX
CC Disclosed peptide, contg. a peptide derived from the human
CC cartilage gp-39 (HC gp 39) autoantigen, can be used to
CC induce specific T cell tolerance to HC gp 39 in a patient with
CC T cell mediated cartilage destruction, esp. rheumatoid arthritis
CC (RA). It may also be used to induce RA in mice, etc. (human disease
CC model). and to diagnose the presence of activated T cells
CC reactive with HC gp 39, i.e. to identify patients suitable for
CC specific treatment.
CC Intraperitoneally the peptide is given in a dose of 0.1 to given/kg
CC i.v. injection, of 10-50 micron/kg to produce a disease model. The
CC peptide has a specific tolerising effect on autoreactive T cells,
CC but leaves other components of the immune system intact, and
CC therefore should have few side effects.
XX
SQ Sequence 13 AA;

Query Match 63.1%; Score 41; DB 12; Length 13;
Best Local Similarity 100.0%; Pred. No. 0.31;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 KSEILASSE 13
Db 1 KSEILASSE 9

RESULT 8
AAW61597
ID AAW61597 standard; protein; 13 AA.
XX
AC AAW61597;
XX
DT 27-MAY 1999 (first entry)
XX
DE Human cartilage gp-39 epitope (1).
XX
KW Human cartilage gp-39; monoclonal antibody; cell surface antigen;
KW rheumatoid arthritis; T cell receptor.
XX
OS Homo sapiens.
XX
PN EP0774261 A1.
XX
PD 21-MAY-1997.
XX
PF 02-DEC-1997; 97EP-0203769.
XX
PR 27-JUN-1997; 97EP-0201972.
PR 06-DEC-1996; 96EP-0203465.
XX
PA (ALKU) AKZO NOBEL NV.
PA (ALKU) AKZO NV.
XX
PI Stoenbakkers PJA;
XX
DR WPI; 1998-400866/35.
XX
PT Preparation of monoclonal antibodies against cell surface antigens
PT used as diagnostic agents and to treat rheumatoid arthritis
PT comprises immunising, then enriching, enriching and cloning B cells
PT and then selecting and cloning hybridoma
XX
PS Examples; Page 9; 23pp; English.
XX
CC The human cartilage gp-39 peptides AAW61597 W61599 were used in the
CC preparation of a monoclonal antibody against a cell surface antigen.
CC The antibody can be used for treatment of rheumatoid arthritis. It
CC is also useful against cell surface antigens and antibodies reactive
CC with the clonotypic structure of T-cell receptors are used as diagnostic
CC reagents. The process allows antibodies to be raised against
CC antigens which are present in very low amounts or which are
CC conformationally unstable without the need to screen very large numbers
CC of hybridomas and without requiring high purity Antigen for enrichment.
XX
SQ Sequence 13 AA;

Query Match 63.1%; Score 41; DB 19; Length 13;
Best Local Similarity 100.0%; Pred. No. 0.31;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 KSEILASSE 13
Db 1 KSEILASSE 9

RESULT 8
AAY69231
ID AAY69231 standard; peptide; 13 AA.
XX
AC AAY69231;
XX
DT 30-MAY 2000 (first entry)
XX
DE Peptide fragment of human cartilage gp-39 (HC gp 39) protein.
XX
KW Human cartilage gp-39 protein; HC gp 39 protein; T cell activator;
KW inflammatory disease; lymphocyte; antigen; bystander suppression;
KW antigen-specific T cell tolerance; autoimmune disorder;
KW inflammatory disease; rheumatoid arthritis;
KW autoimmune inflammatory disease; autoantigen.
XX
OS Homo sapiens.
XX
PN WO200004917-A2.
XX
PD 03-FEB-2000.
XX
PF 19-JUL 1999; 99WO-EP05331.
XX
PR 23-JUL 1998; 98EP-0202471.
XX
PA (ALKU) AKZO NOBEL NV.
XX
PI Miltenburg AMM, Boots AMH;
XX
DR WPI; 2000-182537/16.

```
XX
PI   Use of cartilage protein or peptides for treating inflammatory diseases
PI   especially rheumatoid arthritis -
XX
PS   Claim 2; Page 17; 20pp; English.
XX
CC   AAY77399-403 represent fragments of the human cartilage gp-39 (HC gp-39)
CC   protein. The peptides are T-cell activators. The peptides are used for
CC   preventing inflammatory diseases, by modulating lymphocytes reactive
CC   to antigens other than HC gp-39, present in the same tissue as
CC   HC gp-39, by the induction of antigen specific T cell tolerance.
CC   autoimmune disorders can be treated by bystander suppression. The
CC   HC gp-39 peptides are useful for preventing inflammatory diseases
CC   especially rheumatoid arthritis. The fragments may also be used for
CC   treating autoimmune inflammatory diseases. The HC gp-39 peptides are
CC   especially beneficial in autoimmune conditions in which the
CC   autoantigens associated with the condition are unknown.
XX
SQ   Sequence   13 AA;


  Query Match             63.1%;  Score 41;  DB 21;  Length 13;
  Best Local Similarity   100.0%;  Pred. No. 0.31;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       5 RSPTLASSE 13
           |||||||||
Db       1 RSPTLASSE 9


RESULT   9
AAY77401
ID   AAY77401 standard; peptide; 13 AA.
XX
AC   AAY77401;
XX
DT   22-MAY-2000  (first entry)
XX
DE   Human cartilage glycoprotein-39 (HC gp-39), residues 263-275.
XX
KW   Human cartilage glycoprotein-39; HC gp-39; YKL-40; gp-39;
KW   autoreactive T-cell; tolerisation; desensitisation; autoantigen;
KW   rheumatoid arthritis; human chondrocyte protein; YKL-39;
KW   bystander suppression; autoimmune disorder; Graves' disease;
KW   primary glomerulonephritis; inflammatory bowel disease; diagnostic.
XX
OS   Homo sapiens.
XX
PN   WO200129254-A2.
XX
PD   26-APR-2001.
XX
PF   16-OCT-1999;  99WO-EP05050.
XX
PR   23-OCT-1998;  98EP-0202470.
XX
PA   (ALKE ) AKZO NOBEL NV.
XX
PI   Verbeeten WM;  Boots AMH;
XX
DR   WPI; 2001-367293/38.
XX
PT   New T cell epitope peptides useful for inducing specific T cell
PT   tolerance to an autoantigen for treating autoimmune diseases,
PT   especially rheumatoid arthritis
XX
PS   Example 1; Page 10; 20pp; English.
XX
CC   The invention relates to peptides constituting amino acids 263-275 and
CC   amino acids 266-278 of human chondrocyte protein YKL-39
CC   (AAY77399 Y77400, respectively). The peptides of the invention may be
CC   used in pharmaceutical compositions for the induction of specific T-cell
CC   tolerance (desensitisation) to an autoantigen in patients
CC   suffering from autoimmune disorders, especially rheumatoid arthritis. A
CC   frequently found autoantigen in rheumatoid arthritis sufferers is human
CC   cartilage glycoprotein-39 (HC gp-39, also known as YKL-40). Several
CC   epitopes of HC gp-39 are recognised by peripheral blood T cells from
CC   rheumatoid arthritis patients; one such epitope, HC gp-39 residues
CC   263-275, was recognised by T-cells from 8 out of 9 patients tested.
CC   YKL-39 shares significant sequence identity to HC gp-39. YKL-39 residues
CC   266-278 (AAY77400) constitutes a minimum epitope of HC gp-39 (263-275),
CC   and may be used for desensitisation of T-cells with reactivity to HC
CC   gp-39 (263-275) in rheumatoid arthritis patients. Rheumatoid arthritis
CC   may be treated by the systemic administration of a composition comprising
CC   the peptides of the invention, optionally in combination with existing
CC   treatments. The peptides may also be used for modulating lymphocytes
CC   that are reactive to autoantigens other than YKL-39 and HC gp-39 but
CC   which are present in the same tissue as these autoantigens. By the
CC   induction of antigen-specific T cell tolerance, autoimmune disorders can
CC   be treated by bystander suppression. The conditions that may be treated
CC   via this approach include Graves' disease, primary glomerulonephritis
CC   and inflammatory bowel disease. The peptides, alone with a detection
CC   agent, may additionally be used in a diagnostic composition. The peptides
CC   of the invention have a specific effect on autoreactive T cells whilst
CC   leaving other components of the immune system intact. Hence the
CC   peptides should be safe, having no side effects. The present sequence
CC   represents residues 263-275 of human cartilage glycoprotein-39 (HC gp-39)
CC   which was used in an exemplification of the present invention.
XX
SQ   Sequence   13 AA;


  Query Match             63.1%;  Score 41;  DB 21;  Length 13;
  Best Local Similarity   100.0%;  Pred. No. 0.31;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       5 RSPTLASSE 13
           |||||||||
Db       1 RSPTLASSE 9


RESULT   10
AAU03045
ID   AAU03045 standard; peptide; 13 AA.
XX
AC   AAU03045;
XX
DT   12-SEP-2001  (first entry)
XX
DE   Human cartilage gp-39 autoimmune peptide #1.
XX
KW   Human; autoimmune disease; epitope; antigen; rheumatoid arthritis;
KW   immunosuppression.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT                   /note= "deaminated"
XX
PN   WO200129081-A1.
XX
PD   26-APR-2001.
XX
PF   12-OCT-2000; 2000WO-EP10230.
XX
PR   18-OCT-1999;  99EP-0203427.
XX
PA   (ALKE ) AKZO NOBEL NV.
XX
PI   Van Staveren CJ,  Timmers CM,  Van Galen PJM,  Knoerzer PVA,  Boots AMH;
PI   Miltenburg AMM;
XX
DR   WPI; 2001-367293/38.
XX
PT   Modified peptides derived from human cartilage gp-39 protein epitope
```

```
FT   are useful in immunotherapy, especially in the treatment of arthritis
FT
XX
PS   Claim 12; Page 39; 52pp; English.
XX
CC   The sequence represents Human cartilage gp 39 protein epitope (amino
CC   acid 263-275) which is used as the template for making modified
CC   peptides for use in immunotherapy. The peptides are useful inducing
CC   tolerance induction in patients suffering from autoimmune diseases such
CC   as rheumatoid arthritis. The peptides provide antigen specific, non toxic
CC   immunosuppression. Therapy as a very attractive alternative to
CC   non specific immunosuppression.
XX
SQ   Sequence   13 AA;


  Query Match             63.1%;  Score 41;  DB 22;  Length 13;
  Best Local Similarity   100.0%;  Pred. No. 0.41;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       5 RSFILASSE 13
           |||||||||
Db       3 rsfilasse 9


RESULT  11
AAB03049
ID   AAB03049 standard; peptide; 13 AA.
XX
AC   AAB03049;
XX
DT   12-SEP-2001  (first entry)
XX
DE   Human cartilage gp 39 autoimmune peptide #4.
XX
KW   Human; autoimmune disease; epitope; antigen; rheumatoid arthritis;
KW   immunosuppression.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT                   /note= "N terminal D-1-glutcityl"
XX
PN   WO200129081-A1.
XX
PD   26-APR-2001.
XX
PF   12-OCT-2000; 2000WO-EP10250.
XX
PR   18-OCT-1999;  99EP-0203427.
XX
PA   (AKZO ) AKZO NOBEL NV.
XX
PI   Van Staveren CJ, Timmers CM, Van Galen PJM, Knegt-, RMA  Boots AMH;
PI   Miltenburg AMM;
XX
DR   WPI, 2001-307219/38.
XX
PT   Modified peptides derived from human cartilage gp 39 protein epitope
PT   are useful in immunotherapy, especially in the treatment of arthritis
PT   -
XX
PS   Claim 12; Page 39; 52pp; English.
XX
CC   The sequence represents a modified peptide derived from Human cartilage
CC   gp 39 protein epitope (amino acids 263-275) for use in immunotherapy.
CC   The peptides are useful for inducing tolerance induction in patients
CC   suffering from autoimmune diseases such as rheumatoid arthritis.
CC   The peptides provide antigen specific, non toxic immunosuppression
CC   therapy as a very attractive alternative to non specific
CC   immunosuppression.
XX
SQ   Sequence   13 AA;


  Query Match             63.1%;  Score 41;  DB 22;  Length 13;
  Best Local Similarity   100.0%;  Pred. No. 0.41;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       5 RSFILASSE 13
           |||||||||
```

```
Db      1 rsttlasse 9


RESULT  13
AAU03064
ID   AAU03064 standard; peptide; 13 AA.
XX
AC   AAU03064;
XX
DT   12-SEP 2001  (first entry)
XX
DE   Human cartilage qp-39 autoimmune peptide #19.
XX
KW   Human; autoimmune disease; epitope; antigen; rheumatoid arthritis;
KW   immunosuppression.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified site   1
FT                   /note "N terminal (N-methyl nicotinoyl)"
XX
PN   WO200129081-A1.
XX
PD   26-APR 2001.
XX
PF   12-OCT 2000; 2000WO-EP10230.
XX
PR   18-OCT-1999;   99EP-0204427.
XX
PA   (ALKO ) AKZO NOBEL NV.
XX
PI   Van Staveren CJ, Timmers CM, Van Galen EJM, Kneute, BMA, Boots AMB;
PI   Miltenburg AMM;
XX
DR   WPI; 2001-367294/38.
XX
PT   Modified peptides derived from human cartilage qp-39 protein epitope
PT   are useful in immunotherapy, especially in the treatment of arthritis
PT   -
XX
PS   Claim 12; Page 89; 52pp; English.
XX
CC   The sequence represents a modified peptide derived from Human cartilage
CC   qp-39 protein epitope (amino acids 263-275) for use in immunotherapy.
CC   The peptides are useful for inducing tolerance induction in patients
CC   suffering from autoimmune diseases such as rheumatoid arthritis.
CC   The peptides provide antigen-specific, non-toxic immunosuppression
CC   therapy as a very attractive alternative to non-specific
CC   immunosuppression.
XX
SQ   Sequence   13 AA;


  Query Match             63.1%;  Score 41;  DB 22;  Length 13;
  Best Local Similarity   100.0%;  Pred. No. 0.41;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       5 RSFTLASSE 13
           |||||||||
Db       1 rsttlasse 9


RESULT  14
AAU03063
ID   AAU03063 standard; peptide; 13 AA.
XX
AC   AAU03063;
XX
DT   12-SEP 2001  (first entry)
XX
DE   Human cartilage qp-39 autoimmune peptide #19.
XX
KW   Human; autoimmune disease; epitope; antigen; rheumatoid arthritis;
KW   immunosuppression.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified site   1
FT                   /note "N terminal acetyl"
FT   Modified-site   11
FT                   /label =HER
FT                   /note "Other NH-CH2-CH2-"
FT   Modified-site   13
FT                   /note "C terminus is deleted"
XX
PN   WO200129081-A1.
XX
PD   26-APR 2001.
XX
PF   12-OCT 2000; 2000WO-EP10230.
XX
PR   18-OCT-1999;   99EP-0204427.
XX
PA   (ALKO ) AKZO NOBEL NV.
XX
PI   Van Staveren CJ, Timmers CM, Van Galen EJM, Kneute, BMA, Boots AMB;
PI   Miltenburg AMM;
XX
DR   WPI; 2001-367294/38.
XX
PT   Modified peptides derived from human cartilage qp-39 protein epitope
PT   are useful in immunotherapy, especially in the treatment of arthritis
PT   -
XX
PS   Disclosure; Page 49; 52pp; English.
XX
CC   The sequence represents a modified peptide derived from Human cartilage
CC   qp-39 protein epitope (amino acids 263-275) for use in immunotherapy.
CC   The peptides are useful for inducing tolerance induction in patients
CC   suffering from autoimmune diseases such as rheumatoid arthritis.
CC   The peptides provide antigen-specific, non-toxic immunosuppression
CC   therapy as a very attractive alternative to non-specific
CC   immunosuppression.
XX
SQ   Sequence   13 AA;


  Query Match             63.1%;  Score 41;  DB 22;  Length 13;
  Best Local Similarity   100.0%;  Pred. No. 0.41;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       5 RSFTLASSE 13
           |||||||||
Db       1 rsttlasse 9


RESULT  15
AAU03065
ID   AAU03065 standard; peptide; 13 AA.
XX
AC   AAU03065;
XX
DT   12-SEP 2001  (first entry)
XX
DE   Human cartilage qp-39 autoimmune peptide #21.
XX
KW   Human; autoimmune disease; epitope; antigen; rheumatoid arthritis;
KW   immunosuppression.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
```

```
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT                   /note= "N terminal acetyl"
FT   Misc-Difference 1
FT                   /note= "D-from residue"
XX
PN   W... [illegible] A1.
XX
PD   26-AUG-2001.
XX
PF   12-OCT-2000; 2000WO-EP10230.
XX
PR   14-OCT-1999;  99EP-0203427.
XX
PA   (AUTO) AKZO N.V.
XX
PI   Van Eden W.,  Immers J.M.,  Van Eden G.M.,  Kraan M.A.,  Boots A.M.;
PI   Metselaar J.M.;
XX
DR   WPI; [illegible]/48.
XX
PT   Modified peptides derived from human cartilage gp 39 protein epitope
PT   are useful in immunotherapy, especially in the treatment of arthritis
PT
XX
PS   Disclosure; Page 44; 52pp; English.
XX
CC   The sequence represents a modified peptide derived from human cartilage
CC   gp 39 protein epitope (amino acids 264-275) for use in immunotherapy.
CC   The peptides are useful for inducing tolerance induction in patients
CC   suffering from autoimmune diseases such as rheumatoid arthritis.
CC   The peptides provide antigen specific, non toxic immunosuppression.
CC   therapy as a very attractive alternative to non specific
CC   immunosuppression.
XX
SQ   Sequence  15 AA;
```

```
  Query Match          64.1%;  Score 41;  DB 2;  Length 15;
  Best Local Similarity   100.0%;  Pred. No. 0.61;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          [illegible] 15

Db          [illegible] 9


Search completed: August 26, 2002, 06:16:29
[illegible] DB 2
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 26, 2002, 06:11:33 ; Search Time 5.5 Seconds
(without alignments)
10.504 Million cell updates/sec

Title: US-09-744-282-5
Perfect score: 65
Sequence: 1 PTPGRSFTLASSE 13

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2810428 seqs, 244255594 residues

Total number of hits satisfying chosen parameters: 88388

Minimum DB seq length: 0
Maximum DB seq length: 13

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
[illegible]

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	65	100.0	13	1	US-08-619-645-5	Sequence 5, Appli
2	65	100.0	13	2	US-08-634-493-5	Sequence 5, Appli
3	44	67.7	9	1	US-08-619-645-1	Sequence 1, Appli
4	44	67.7	9	2	US-08-634-493-1	Sequence 1, Appli
5	41	63.1	13	1	US-08-619-645-6	Sequence 6, Appli
6	41	63.1	13	2	US-08-634-493-6	Sequence 6, Appli
7	40	61.5	9	1	US-08-619-645-2	Sequence 2, Appli
8	40	61.5	9	2	US-08-634-493-2	Sequence 2, Appli
9	34	52.3	9	4	US-09-042-358-331	Sequence 331, App
10	34	52.3	9	4	US-08-758-417A-179	Sequence 179, App
11	34	52.3	10	1	US-08-154-829-41	Sequence 41, Appl
12	34	52.3	10	1	US-08-584-799-41	Sequence 41, Appl
13	34	52.3	10	3	US-08-159-339A-351	Sequence 351, App
14	34	52.3	10	3	US-08-159-339A-353	Sequence 353, App
15	34	52.3	12	1	US-07-778-238B-21	Sequence 21, Appl
16	34	52.3	12	1	US-07-963-321-21	Sequence 21, Appl
17	34	52.3	12	1	US-08-290-641-21	Sequence 21, Appl
18	34	52.3	12	1	US-08-548-543-21	Sequence 21, Appl
19	34	52.3	12	5	PCT-US96-09859-21	Sequence 21, Appl
20	34	52.3	13	1	US-08-649-272A-8	Sequence 8, Appli
21	34	52.3	13	3	US-09-105-799-8	Sequence 8, Appli
22	33	50.8	8	4	US-08-590-563-xx	Sequence xx, Appl
23	33	50.8	9	4	US-09-187-859-3857	Sequence 3857, Ap
24	33	50.8	10	5	PCT-US92-07865-11	Sequence 11, Appl
25	33	50.8	7	2	US-09-070-756-11	Sequence 11, Appl
26	33	50.8	7	2	US-09-070-756-12	Sequence 12, Appl
27	33	50.8	9	2	US-08-340-283-102	Sequence 102, App
28	33	50.8	9	3	US-08-159-339A-334	Sequence 334, App
29	33	50.8	9	3	US-08-159-339A-366	Sequence 366, App
30	33	50.8	9	3	US-08-159-339A-394	Sequence 394, App
31	33	50.8	9	3	US-08-159-339A-460	Sequence 460, App
32	33	50.8	9	3	US-08-159-339A-434	Sequence 434, App
33	33	50.8	9	3	US-08-159-339A-444	Sequence 444, App
34	33	50.8	9	4	US-09-105-830-25	Sequence 25, Appl
35	33	50.8	9	4	US-09-105-830-55	Sequence 55, Appl
36	33	50.8	9	4	US-09-105-830-67	Sequence 67, Appl
37	33	50.8	9	5	PCT-US95-16415-84	Sequence 84, Appl
38	33	50.8	10	2	US-08-747-285-18	Sequence 18, Appl
39	33	50.8	10	2	US-08-747-285A-16	Sequence 16, Appl
40	33	50.8	10	2	US-08-685-589A-102	Sequence 102, App
41	33	50.8	10	3	US-08-159-339A-375	Sequence 375, App
42	33	50.8	10	3	US-08-159-339A-452	Sequence 452, App
43	33	50.8	10	4	US-09-105-830-29	Sequence 29, Appl
44	33	50.8	10	4	US-09-105-830-72	Sequence 72, Appl
45	33	50.8	12	1	US-08-260-582-18	Sequence 18, Appl

ALIGNMENTS

RESULT 1
US-08-619-645-5
; Sequence 5, Application US/08/619,645
; Patent No. 5736507
; GENERAL INFORMATION:
; APPLICANT: Boots, Anna M.B.
; APPLICANT: Verheijden, Gijsbertus J.M.
; TITLE OF INVENTION: [illegible]
; TITLE OF INVENTION: Antagonist for use in immunotherapy of autoimmune
; TITLE OF INVENTION: disease
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Akzo No. 5736507el Patent Department
; STREET: 1300 Piccard Drive, Suite 206
; CITY: Rockville
; STATE: Maryland
; COUNTRY: U.S.A.
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/619,645
; FILING DATE: 25-MAR-1996
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Gormley, Mary E.
; REGISTRATION NUMBER: 34,409
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 947-4433
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 13 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-619-645-5

Query Match 100.0%; Score 65; DB 1; Length 13;
Best Local Similarity 100.0%; Pred. No. 1.9e-06;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PTPGRSFTLASSE 13
||||||||||||| |
Db 1 PTPGRSFTLASSE 13

```
RESULT   3
PCT/EP94/04201
; Sequence 1, Application US/08634493
; Patent No. 5843449
; GENERAL INFORMATION:
;     APPLICANT:  A.M.E. Boots
;     APPLICANT:  G.F.M. Verheijden
;     APPLICANT:  E.S. Bos
;     TITLE OF INVENTION:  No. 5843449-1 Peptides derived from autoantigen for use
;     TITLE OF INVENTION:  in Immunotherapy of Autoimmune Diseases
;     NUMBER OF SEQUENCES:  10
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Akzo No. 5843449-1 Patent Department
;       STREET:  1300 Piccard Drive, Suite 206
;       CITY:  Rockville
;       STATE:  Maryland
;       COUNTRY:  USA
;       ZIP:  20850
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC DOS/MS DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/634,493
;       FILING DATE:  18-APR-1996
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/619,645
;       FILING DATE:  25-MAR-1996
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  PCT/EP95/04201
;       FILING DATE:  25-OCT-1995
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  NL 9420 1287
;       ADDRESSEE:  Akzo No. 5736507-1 Patent Department
;       STREET:  1300 Piccard Drive, Suite 206
;       CITY:  Rockville
;       STATE:  Maryland
;       COUNTRY:  U.S.A.
;       ZIP:  20850
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC DOS/MS DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25 (EPO)
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/619,645
;       FILING DATE:  25-MAR-1996
;       CLASSIFICATION:  514
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Gormley, Mary E.
;       REGISTRATION NUMBER:  34,409
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (301) 947-4433
;   INFORMATION FOR SEQ ID NO:  1:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  9 amino acids
;       TYPE:  amino acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  peptide
US-08-619-645-1


  Query Match             67.7%;  Score 44;  DB 1;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 1.7e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3 FGRSFILAS 11
            |||||||||
Db       1 FGRSFILAS 9


RESULT   4
US-08-634-493-1
; Sequence 1, Application US/08634493
; Patent No. 5843449
; GENERAL INFORMATION:
;     APPLICANT:  A.M.E. Boots
;     APPLICANT:  G.F.M. Verheijden
;     APPLICANT:  E.S. Bos
;     TITLE OF INVENTION:  No. 5843449-1 Peptides derived from autoantigen for use
;     TITLE OF INVENTION:  in Immunotherapy of Autoimmune Diseases
;     NUMBER OF SEQUENCES:  10
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Akzo No. 5843449-1 Patent Department
;       STREET:  1300 Piccard Drive, Suite 206
;       CITY:  Rockville
;       STATE:  Maryland
;       COUNTRY:  USA
;       ZIP:  20850
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC DOS/MS DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/634,493
;       FILING DATE:  18-APR-1996
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/619,645
;       FILING DATE:  25-MAR-1996
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  PCT/EP95/04201
;       FILING DATE:  25-OCT-1995
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  NL 9420 1287
```

```
;       FILING DATE:  27-OCT-1994
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  NL 9520088600
;       FILING DATE:  07-APR-1995
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (301) 258-5200
;       TELEFAX:  (301) 977-0847
;   INFORMATION FOR SEQ ID NO:  1:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  9 amino acids
;       TYPE:  amino acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  peptide
US-08-634-493-1


  Query Match             67.7%;  Score 44;  DB 2;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 1.7e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       5 RSFTLASSE 13
           |||||||||
Db       1 RSFTLASSE 9


RESULT   5
US-08-619-645-5
; Sequence 5, Application US/08619645
; Patent No. 5736507
;  GENERAL INFORMATION:
;    APPLICANT:  Boots, Anna M.H.
;    APPLICANT:  Verheijden, Gijsbertus F.M.
;    TITLE OF INVENTION:  No. 5736507el peptides derived from
;    TITLE OF INVENTION:  autoantigen for use in immunotherapy of autoimmune
;    TITLE OF INVENTION:  disease
;    NUMBER OF SEQUENCES:  9
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Akzo No. 5736507el Patent Department
;      STREET:  1300 Piccard Drive, Suite 206
;      CITY:  Rockville
;      STATE:  Maryland
;      COUNTRY:  U.S.A.
;      ZIP:  20850
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25(EPO)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/619,645
;      FILING DATE:  25-MAR-1996
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Gormley, Mary E.
;      REGISTRATION NUMBER:  34,409
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (301) 947-4433
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  13 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-619-645-5


  Query Match             63.1%;  Score 41;  DB 1;  Length 13;
  Best Local Similarity   100.0%;  Pred. No. 0.073;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       5 RSFTLASSE 13
           |||||||||
Db       1 RSFTLASSE 9


RESULT   6
US-08-634-493-6
; Sequence 6, Application US/08634493
; Patent No. 5843449
;  GENERAL INFORMATION:
;    APPLICANT:  A.M.H. Boots
;    APPLICANT:  G.F.M. Verheijden
;    APPLICANT:  E.S. Bos
;    TITLE OF INVENTION:  No. 5843449el Peptides derived from autoantigen for use
;    TITLE OF INVENTION:  in Immunotherapy of Autoimmune Diseases
;    NUMBER OF SEQUENCES:  10
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Akzo No. 5843449el Patent Department
;      STREET:  1300 Piccard Drive, Suite 206
;      CITY:  Rockville
;      STATE:  Maryland
;      COUNTRY:  USA
;      ZIP:  20850
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/634,493
;      FILING DATE:  18-APR-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/619,645
;      FILING DATE:  25-MAR-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/EP95/04201
;      FILING DATE:  25-OCT-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  NL 9420001287
;      FILING DATE:  27-OCT-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  NL 9520088600
;      FILING DATE:  07-APR-1995
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (301) 258-5200
;      TELEFAX:  (301) 977-0847
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  13 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-634-493-6


  Query Match             63.1%;  Score 41;  DB 2;  Length 13;
  Best Local Similarity   100.0%;  Pred. No. .07;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       5 RSFTLASSE 13
           |||||||||
Db       1 RSFTLASSE 9


RESULT   7
US-08-619-645-2
; Sequence 2, Application US/08619645
; Patent No. 5736507
;  GENERAL INFORMATION:
;    APPLICANT:  Boots, Anna M.H.
;    APPLICANT:  Verheijden, Gijsbertus F.M.
;    TITLE OF INVENTION:  No. 5736507el peptides derived from
;    TITLE OF INVENTION:  autoantigen for use in immunotherapy of autoimmune
```

; TITLE OF INVENTION: disease
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Akzo No. 5746507el Patent Department
; STREET: 1300 Piccard Drive, Suite 206
; CITY: Rockville
; STATE: Maryland
; COUNTRY: U.S.A.
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC DOS/MS DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/619,645
; FILING DATE: 25-MAR-1996
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Gormley, Mary E.
; REGISTRATION NUMBER: 34,409
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 947-4433
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-619-645-2

Query Match 49.2%; Score 42; DB 2; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.7e-05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7 FILASSE 13
||||||
Db 1 FILASSE 7

RESULT 9
US-09-042-353-331
; Sequence 331, Application US/09042353
; Patent No. 6255458
; GENERAL INFORMATION:
; APPLICANT: Lonberg, Nils
; APPLICANT: Kay, Robert M.
; TITLE OF INVENTION: Transgenic No. 6255458 Human Animals for
; TITLE OF INVENTION: Producing Heterologous Antibodies
; NUMBER OF SEQUENCES: 421
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/042,353
; FILING DATE: 13-MAR-1998
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/810,279
; FILING DATE: 17-DEC-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/853,408
; FILING DATE: 18-MAR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/904,068
; FILING DATE: 23-JUN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/990,860
; FILING DATE: 16-DEC-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/053,131
; FILING DATE: 26-APR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/096,762
; FILING DATE: 22-JUL-1993

```
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/175,381
;       FILING DATE:  28-NOV-1993
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/161,739
;       FILING DATE:  03-DEC-1993
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/145,699
;       FILING DATE:  30-DEC-1993
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/209,741
;       FILING DATE:  09-MAR-1994
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/352,422
;       FILING DATE:  07-DEC-1994
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/541,181
;       FILING DATE:  10-OCT-1995
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/728,464
;       FILING DATE:  10-OCT-1996
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  WO PCT/US96/16434
;       FILING DATE:  10-OCT-1996
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/758,417
;       FILING DATE:  02-DEC-1996
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  WO PCT/US97/21803
;       FILING DATE:  01-DEC-1997
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Apple, Randolph T.
;       REGISTRATION NUMBER:  36,429
;       REFERENCE/DOCKET NUMBER:  018443-000500US
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (415) 576-0200
;       TELEFAX:  (415) 576-0300
;   INFORMATION FOR SEQ ID NO:  331:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  9 amino acids
;       TYPE:  amino acid
;       STRANDEDNESS:
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  peptide
US-09-342-353-331


  Query Match             36.9%;  Score 24;  DB 4;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 1.7e+05;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 PTFG 4
            ||||
Db       6 PTFG 9


RESULT  10
US-08-758-417A-179
; Sequence 179, Application US/08758417A
; Patent No. 5850229
;   GENERAL INFORMATION:
;        APPLICANT: Lerdrup, Nils
;                   Kay, Robert M.
;        TITLE OF INVENTION: Transgenic Non-Human Animals for
;                            Producing Heterologous Antibodies
;        NUMBER OF SEQUENCES: 417
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Townsend and Townsend and Crew LLP
;             STREET: Two Embarcadero Center, Eighth Floor
;             CITY: San Francisco
;             STATE: California
;             COUNTRY: USA
;             ZIP: 94111-3834
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.3
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/08/758,417A
;             FILING DATE: 02-Dec-1996
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: US 08/728,464
;             FILING DATE: 10-OCT-1996
;             APPLICATION NUMBER: US 08/544,404
;             FILING DATE: 10-OCT-1995
;             APPLICATION NUMBER: US 08/352,422
;             FILING DATE: 07-DEC-1994
;             APPLICATION NUMBER: US 08/209,741
;             FILING DATE: 09-MAR-1994
;             APPLICATION NUMBER: US 08/161,739
;             FILING DATE: 10-DEC-1993
;             APPLICATION NUMBER: US 08/161,739
;             FILING DATE: 03-DEC-1993
;             APPLICATION NUMBER: US 08/145,699
;             FILING DATE: 18-NOV-1993
;             APPLICATION NUMBER: US 08/096,762
;             FILING DATE: 22-JUL-1993
;             APPLICATION NUMBER: US 08/053,191
;             FILING DATE: 26-APR-1993
;             APPLICATION NUMBER: US 07/990,860
;             FILING DATE: 16-DEC-1992
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Serafini, Andrew T.
;             REGISTRATION NUMBER: 41,303
;             REFERENCE/DOCKET NUMBER: 018443-000500US
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (415) 576-0200
;             TELEFAX: (415) 576-0300
;   INFORMATION FOR SEQ ID NO: 179:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 9 amino acids
;             TYPE: amino acid
;             STRANDEDNESS: <Unknown>
;             TOPOLOGY: linear
;        MOLECULE TYPE: peptide
;        SEQUENCE DESCRIPTION: SEQ ID NO: 179:
US-08-758-417A-179


  Query Match             36.9%;  Score 24;  DB 4;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 1.7e+05;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 PTFG 4
            ||||
Db       6 PTFG 9


RESULT  11
US-08-164-839-41
; Sequence 41, Application US/08164839
; Patent No. 5514573
;  GENERAL INFORMATION:
;    APPLICANT:  YASUEDA, HISASHI
;    APPLICANT:  NAKANISHI, KAZUO
;    APPLICANT:  MOTOKI, MASAO
;    APPLICANT:  NAGASE, KAZUO
;    APPLICANT:  MATSUI, HIROSHI
;    TITLE OF INVENTION:  GENE ENCODING GLUTAMINASE DERIVED
;    TITLE OF INVENTION:  FROM FISH
;    NUMBER OF SEQUENCES:  72
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  OBLON, SPIVAK, McCLELLAND, MAIER & NEUSTADT
;      ADDRESSEE:  P.C.
```

```
      ;  STREET:  1755 Jefferson Davis Highway, Fourth Floor
      ;  CITY:  Arlington
      ;  STATE:  Virginia
      ;  COUNTRY:  U.S.A.
      ;  ZIP:  22202
      ;  COMPUTER READABLE FORM:
      ;  MEDIUM TYPE:  Floppy disk
      ;  COMPUTER:  IBM PC compatible
      ;  OPERATING SYSTEM:  PC DOS/MS DOS
      ;  SOFTWARE:  PatentIn Release #1.0, Version #1.3
      ;  CURRENT APPLICATION DATA:
      ;  APPLICATION NUMBER:  08/584,799
      ;  FILING DATE:
      ;  CLASSIFICATION:  435
      ;  PRIOR APPLICATION DATA:
      ;  APPLICATION NUMBER:  08/004,729
      ;  FILING DATE:  14-JAN-1993
      ;  ATTORNEY/AGENT INFORMATION:
      ;  NAME:  Oblon, No. 560/84Oman F.
      ;  REGISTRATION NUMBER:  24,618
      ;  REFERENCE/DOCKET NUMBER:  10 599-0
      ;  TELECOMMUNICATION INFORMATION:
      ;  TELEPHONE:  (703)412 3000
      ;  TELEFAX:  (703)413-2220
      ;  TELEX:  248855 OPAT UR
      ;  INFORMATION FOR SEQ ID NO:  41:
      ;  SEQUENCE CHARACTERISTICS:
      ;  LENGTH:  10 amino acids
      ;  TYPE:  amino acid
      ;  TOPOLOGY:  linear
      ;  MOLECULE TYPE:  peptide
US 08/584,799-41

  Query Match  96.9%;  Score 24;  DB 1;  Length 10;
  Best Local Similarity  57.1%;  Pred. No. 97;
  Matches  4;  Conservative  0;  Mismatches  3;  Indels  0;  Gaps  0;

Qy  1 PTEGRSE 7
      | || |
Db  1 PVEGIDF 7


RESULT 13
US-08-159,339A-351
; Sequence 351, Application US/08159339A
; Patent No. 6037135
; GENERAL INFORMATION:
;  APPLICANT:  Kubo, Ralph T.
;  APPLICANT:  Grey, Howard M.
;  APPLICANT:  Sette, Alessandro
;  APPLICANT:  Celis, Esteban
;  TITLE OF INVENTION:  HLA Binding peptides and Their
;  TITLE OF INVENTION:  Uses
;  NUMBER OF SEQUENCES:  1254
;  CORRESPONDENCE ADDRESS:
;  ADDRESSEE:  Townsend and Townsend and Crew LLP
;  STREET:  Two Embarcadero Center, Eighth Floor
;  CITY:  San Francisco
;  STATE:  CA
;  COUNTRY:  USA
;  ZIP:  94111-3834
;  COMPUTER READABLE FORM:
;  MEDIUM TYPE:  diskette
;  COMPUTER:  IBM Compatible
;  OPERATING SYSTEM:  DOS
;  SOFTWARE:  FastSEQ for Windows Version 2.0
;  CURRENT APPLICATION DATA:
;  APPLICATION NUMBER:  US/08/159,339A
;  FILING DATE:  29-NOV-1993
;  CLASSIFICATION:  424
;  PRIOR APPLICATION DATA:
;  APPLICATION NUMBER:  US 07/926,666
;  FILING DATE:  07-AUG-1992
;  APPLICATION NUMBER:  US 08/027,746
;  FILING DATE:  05 MAR 1993
;  APPLICATION NUMBER:  US 08/116,996
;  FILING DATE:  06 AUG-1993
;  ATTORNEY/AGENT INFORMATION:
;  NAME:  Weber, Ellen Lauver
;  REGISTRATION NUMBER:  32,762
;  REFERENCE/DOCKET NUMBER:  018128.0053/P0S
;  TELECOMMUNICATION INFORMATION:
;  TELEPHONE:  (415) 576-0200
;  TELEFAX:  (415) 576-0300
;  TELEX:
;  INFORMATION FOR SEQ ID NO:  351:
```

; SEQUENCE CHARACTERISTICS:
; LENGTH: 12 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-159-649A-454

Query Match 36.9%; Score 24; DB 3; Length 12;
Best Local Similarity 62.5%; Pred. No. 97;
Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2 TPGRSFTL 9
|||| |
Db 2 TPGRKLHL 9

RESULT 14
US-08-159-649A-453
; Sequence 453, Application US/08159649A
; Patent No. 5677435
; GENERAL INFORMATION:
; APPLICANT: Kubo, Ralph T.
; APPLICANT: Grey, Howard M.
; APPLICANT: Sette, Alessandro
; APPLICANT: Celis, Esteban
; TITLE OF INVENTION: HLA Binding peptides and Their
; TITLE OF INVENTION: Uses
; NUMBER OF SEQUENCES: 1254
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/159,649A
; FILING DATE: 29-NOV-1993
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/926,666
; FILING DATE: 07-AUG-1992
; APPLICATION NUMBER: US 08/027,746
; FILING DATE: 05-MAR-1993
; APPLICATION NUMBER: US 08/103,396
; FILING DATE: 06-AUG-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Weber, Ellen Lauver
; REGISTRATION NUMBER: 32,762
; REFERENCE/DOCKET NUMBER: 019623-005040US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; TELEX:
; INFORMATION FOR SEQ ID NO: 453:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 12 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-159-649A-453

Query Match 36.9%; Score 24; DB 3; Length 12;
Best Local Similarity 62.5%; Pred. No. 97;
Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2 TPGRSFTL 9
|||| |
Db 2 TPGRKLHL 9

RESULT 15
US-07-778-233B-21
; Sequence 21, Application US/07778233B
; Patent No. 5270170
; GENERAL INFORMATION:
; APPLICANT: Schatz, Peter J.
; APPLICANT: Cull, Millard G.
; APPLICANT: Miller, Jeff F.
; TITLE OF INVENTION: Peptide Library and Screening Method
; NUMBER OF SEQUENCES: 78
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: William M. Smith
; STREET: One Market Plaza, Steuart Tower, Suite 1000
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/778,233B
; FILING DATE: 19911016
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith, William M.
; REGISTRATION NUMBER: 30,223
; REFERENCE/DOCKET NUMBER: 11509-50
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-326-2400
; TELEFAX: 415-326-2422
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 12 amino acids
; TYPE: AMINO ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; IMMEDIATE SOURCE:
; CLONE: 46.3.0.3
US-07-778-233B-21

Query Match 36.9%; Score 24; DB 1; Length 12;
Best Local Similarity 62.5%; Pred. No. 1.2e+02;
Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 4 GRSFTLAS 11
|||| : |
Db 3 GRSFHVIS 10

Search completed: August 26, 2002, 06:15:41
Job time: 128 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein: protein search, using sw model

Run on: August 26, 2002, 06:12:49 ; Search Time 50.45 Seconds
(without alignments)
24.756 Million cell updates/sec

Title: US-09-744-282-6
Perfect score: 62
Sequence: 1 ASFILASSETGVG 13

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 1623

Minimum DB seq length: 0
Maximum DB seq length: 13

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_71:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	28	45.2	7	2	PT0702	T-cell receptor be
2	28	45.2	9	2	PH0935	T-cell receptor be
3	27	43.5	8	2	PT0627	T-cell receptor be
4	25	40.3	6	2	PT0619	T-cell receptor be
5	24	38.7	10	2	PH0900	T cell receptor be
6	23	37.1	8	2	PT0509	T-cell receptor be
7	22	35.5	8	2	PI0631	T cell receptor be
8	22	35.5	8	2	PT0554	T cell receptor be
9	22	35.5	12	2	PH1187	T-cell receptor al
10	21	33.9	6	2	PT0630	T-cell receptor be
11	21	33.9	6	2	PT0668	T cell receptor be
12	21	33.9	7	2	PT0667	T cell receptor be
13	21	33.9	7	2	PT0719	T cell receptor be
14	21	33.9	9	2	PH0918	T-cell receptor be
15	21	33.9	10	2	G60787	sperm activating p
16	21	33.9	10	2	I60588	sperm activating p
17	21	33.9	10	2	PH0946	T cell receptor be
18	21	33.9	11	2	PH0906	T cell receptor be
19	21	33.9	13	2	PQ0491	serine incompatibil
20	21	33.9	13	2	D56661	Sclerus specific a
21	20	32.3	6	2	PT0550	T cell receptor be
22	20	32.3	7	2	PI0524	T cell receptor be
23	20	32.3	7	2	PT0683	T cell receptor be
24	20	32.3	12	2	A28955	polysaccharide prot
25	20	32.3	13	2	B28955	polysaccharide prot
26	20	32.3	13	2	PH1479	T cell receptor be
27	19.5	31.5	7	2	PT0529	T cell receptor be
28	19	30.6	5	2	PI0561	T cell receptor be
29	19	30.6	6	2	PT0643	T cell receptor be
30	19	30.6	8	2	PI0559	T cell receptor be
31	19	30.6	10	2	D60567	serine threonine p
32	19	30.6	10	2	S28471	T cell receptor al
33	19	30.6	11	2	PH0969	T cell receptor be
34	19	30.6	12	4	S49078	frame shifted cyta
35	19	30.6	13	2	A28995	Glycosin heavy chain
36	19	30.6	13	2	PH0928	T cell receptor be
37	18.5	29.8	8	2	PI0595	T cell receptor be
38	18	29.0	5	2	PT0610	T cell receptor be
39	18	29.0	6	2	PI0516	T cell receptor be
40	18	29.0	6	2	PI0637	T cell receptor be
41	18	29.0	6	2	PT0720	T cell receptor be
42	18	29.0	7	2	PI0523	T cell receptor be
43	18	29.0	7	2	PI0642	T cell receptor be
44	18	29.0	8	2	A37521	R-phycoerythrin-sa
45	18	29.0	9	2	PH0902	T cell receptor be

ALIGNMENTS

RESULT 1
PT0762
T-cell receptor beta chain V D-J region (15d 2AA) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 31-Mar-1997
C;Accession: PT0702
R;Feeney, A.J.
J. Exp. Med. 174, 115-124, 1991
A;Title: Junctional sequences of fetal T-cell receptor beta chains have few N-regions
A;Reference number: PI0509; MUID:91377341
A;Accession: PT0702
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-7 <FEE>
A;Experimental source: newborn thymus; strain BALB/C
C;Keywords: T-cell receptor

Query Match 45.2%; Score 28; DB 2; Length 7;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 ASSETG 11
||||||
Db 1 ASSETG 6

RESULT 2
PH0935
T-cell receptor beta chain V D-J region (cl beta-5) - rat (fragment)
C;Species: Rattus norvegicus (Norway rat)
C;Date: 09-Oct-1992 #sequence_revision 09-Oct-1992 #text_change 30-May-1997
C;Accession: PH0935
R;Gold, D.P.; Offner, H.; Sun, D.; Wiley, S.; Vandenbark, A.A.; Wilson, D.B.
J. Exp. Med. 174, 1467-1476, 1991
A;Title: Analysis of T-cell receptor beta chains in Lewis rats with experimental allo
A;Reference number: PH0891; MUID:92078857
A;Accession: PH0935
A;Molecule type: mRNA
A;Residues: 1-9 <GOL>
A;Experimental source: complete Freund's adjuvant-immunized lymph node
C;Keywords: T-cell receptor

Query Match 45.2%; Score 28; DB 2; Length 9;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 ASSETG 11
||||||
Db 2 ASSETG 7

RESULT 5
PT0627
T-cell receptor beta chain V-D-J region (100-2N) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
C;Accession: PT0627
R;Feeney, A.J.
J. Exp. Med. 174, 115-124, 1991
A;Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.
A;Reference number: PT0509; MUID:91277601
A;Accession: PT0627
A;Status: translation not shown
A;Molecule type: mRNA
A;Residues: 1-8 <FEE>
A;Experimental source: newborn thymus, strain BALB/c
C;Keywords: T cell, receptor

Query Match 43.5%; Score 27; DB 2; Length 8;
Best Local Similarity 62.5%; Prod. No. 2.8e-05;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 6 ASSETGVG 13
|||:::|:
Db 2 ASSDAGGG 8

RESULT 4
PT0629
T cell receptor beta chain V-D-J region (120-2N) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
C;Accession: PT0629; PT0563; PT0598
R;Feeney, A.J.
J. Exp. Med. 174, 115-124, 1991
A;Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.
A;Reference number: PT0509; MUID:91277601
A;Accession: PT0629
A;Status: translation not shown
A;Molecule type: mRNA
A;Residues: 1-8 <FEE>
A;Experimental source: newborn thymus, strain BALB/c, clone 120-2N
A;Accession: PT0563
A;Status: translation not shown
A;Molecule type: mRNA
A;Residues: 1-8 <FEE>
A;Experimental source: day 18 fetal thymus, strain BALB/c, clone 26-5A
A;Accession: PT0598
A;Status: translation not shown
A;Molecule type: mRNA
A;Residues: 1-8 <FEE>
A;Experimental source: newborn thymus, strain BALB/c, clone III 12
C;Keywords: T cell receptor

Query Match 43.5%; Score 27; DB 2; Length 8;
Best Local Similarity 62.5%; Prod. No. 2.8e-05;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 6 ASSETGVG 13
|||:::|:
Db 2 ASSDAGGG 8

RESULT 6
PT0509
T-cell receptor beta chain V-D-J region (100-4N) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
C;Accession: PT0509
R;Feeney, A.J.
J. Exp. Med. 174, 115-124, 1991
A;Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.
A;Reference number: PT0509; MUID:91277601
A;Accession: PT0509
A;Status: translation not shown
A;Molecule type: mRNA
A;Residues: 1-8 <FEE>
A;Experimental source: adult thymus, strain BALB/c
C;Keywords: T cell receptor

Query Match 37.1%; Score 23; DB 2; Length 8;
Best Local Similarity 62.5%; Prod. No. 2.8e-05;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 6 ASSETGVG 13
||| : | |
Db 1 ASSDAGGG 8

RESULT 7
PT0631
T cell receptor beta chain V-D-J region (111-11) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
C;Accession: PT0631
R;Feeney, A.J.
J. Exp. Med. 174, 115-124, 1991
A;Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.
A;Reference number: PT0509; MUID:91277601
A;Accession: PT0631
A;Status: translation not shown
A;Molecule type: mRNA
A;Residues: 1-8 <FEE>
A;Experimental source: newborn thymus, strain BALB/c
C;Keywords: T-cell receptor

Query Match 35.5%; Score 22; DB 2; Length 8;
Best Local Similarity 50.0%; Prod. No. 2.8e+00;
Matches 4; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 6 ASSETGVG 13
|| : | |
Db 1 ASGDAGTG 8

RESULT 8
A;Title: Analysis of T cell receptor beta chains in Lewis rats with experimental alle
A;Reference number: PH0891; MUID:92078857
A;Accession: PH0900
A;Molecule type: mRNA
A;Residues: 1-10 <GOL>
A;Experimental source: myelin basic protein-immunized T cell
C;Keywords: T-cell receptor

Query Match 48.7%; Score 24; DB 2; Length 10;
Best Local Similarity 57.1%; Pred. No. 2e+02;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 6 ASSETGV 12
|||:::|:
Db 2 ASSDSGI 8

RESULT 6
PT0509
T-cell receptor beta chain V-D-J region (100-4N) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
C;Accession: PT0509
R;Feeney, A.J.
J. Exp. Med. 174, 115-124, 1991
A;Title: Junctional sequences of fetal T cell receptor beta chains have few N regions
A;Reference number: PT0509; MUID:91277601
A;Accession: PT0509
A;Status: translation not shown
A;Molecule type: mRNA
A;Residues: 1-8 <FEE>
A;Experimental source: adult thymus, strain BALB/c
C;Keywords: T cell receptor

Query Match 37.1%; Score 23; DB 2; Length 8;
Best Local Similarity 62.5%; Prod. No. 2.8e-05;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 6 ASSETGVG 13
||| : | |
Db 1 ASGDAGGG 8

RESULT 7
PT0631
T cell receptor beta chain V-D-J region (111-11) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
C;Accession: PT0631
R;Feeney, A.J.
J. Exp. Med. 174, 115-124, 1991
A;Title: Junctional sequences of fetal T cell receptor beta chains have few N regions
A;Reference number: PT0509; MUID:91277601
A;Accession: PT0631
A;Status: translation not shown
A;Molecule type: mRNA
A;Residues: 1-8 <FEE>
A;Experimental source: newborn thymus, strain BALB/c
C;Keywords: T-cell receptor

Query Match 35.5%; Score 22; DB 2; Length 8;
Best Local Similarity 50.0%; Prod. No. 2.8e+00;
Matches 4; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 6 ASSETGVG 13
|| : | |
Db 1 ASGDAGTG 8

RESULT 8

```
PT0554
T-cell receptor beta chain V-D-J region (126.1G)  mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 9-May-1997
C;Accession: PT0554
R;Feeney, A.J.
J. Exp. Med. 174, 115-124, 1991
A;Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.
A;Reference number: PT0509; MUID:91277601
A;Accession: PT0554
A;Status: translation not shown
A;Molecule type: mRNA
A;Residues: 1-6 <FEE>
A;Experimental source: day 18 fetal thymus, strain BALB/c
C;Keywords: T-cell receptor


  Query Match             35.5%;  Score 22;  DB 2;  Length 6;
  Best Local Similarity   50.0%;  Pred. No. 2.8e-05;
  Matches    4;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       6 ASSETGVG 13
           :| : :|
Db       1 ASGDEGIA 8


RESULT  9
PH1187
T-cell receptor alpha chain V region (CW3/IF11) - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 16-Jul-1999
C;Accession: PH1187
R;Casanova, J.L.; Cerottini, J.C.; Matthes, M.; Necker, A.; Gournier, B.; Barra, C.; Wid
J. Exp. Med. 175, 439-447, 1992
A;Title: H-2-restricted cytolytic T lymphocytes specific for HIV display T cell receptor
A;Reference number: S26512; MUID:92364546
A;Accession: PH1187
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-12 <CAS>


  Query Match             35.5%;  Score 22;  DB 2;  Length 12;
  Best Local Similarity   83.3%;  Pred. No. 5.9e-02;
  Matches    5;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       6 ASSETG 11
           :|||:|
Db       2 AVSETG 7


RESULT  10
PT0630
T-cell receptor beta chain V-D-J region (111-1C)  mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
C;Accession: PT0630
R;Feeney, A.J.
J. Exp. Med. 174, 115-124, 1991
A;Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.
A;Reference number: PT0509; MUID:91277601
A;Accession: PT0630
A;Status: translation not shown
A;Molecule type: mRNA
A;Residues: 1-6 <FEE>
A;Experimental source: newborn thymus, strain BALB/c
C;Keywords: T-cell receptor


  Query Match             33.9%;  Score 21;  DB 2;  Length 6;
  Best Local Similarity   66.7%;  Pred. No. 2.8e-05;
  Matches    4;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       6 ASSETG 11
           |||:||
Db       1 ASGDTG 6


RESULT  11
PT0668
T-cell receptor beta chain V-D-J region (121-4B8)  mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
C;Accession: PT0668
R;Feeney, A.J.
J. Exp. Med. 174, 115-124, 1991
A;Title: Junctional sequences of fetal T cell receptor beta chains have few N regions
A;Reference number: PT0509; MUID:91277601
A;Accession: PT0668
A;Status: translation not shown
A;Molecule type: mRNA
A;Residues: 1-6 <FEE>
A;Experimental source: day 4 [illegible]
C;Keywords: T-cell receptor


  Query Match             33.9%;  Score 21;  DB 2;  Length 6;
  Best Local Similarity   83.3%;  Pred. No. 2.8e-05;
  Matches    5;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       6 ASSETG 11
           |||:|
Db       1 ASSGTG 6


RESULT  12
PT0667
T-cell receptor beta chain V-D-J region (121-2J)  mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
C;Accession: PT0667
R;Feeney, A.J.
J. Exp. Med. 174, 115-124, 1991
A;Title: Junctional sequences of fetal T cell receptor beta chains have few N regions
A;Reference number: PT0509; MUID:91277601
A;Accession: PT0667
A;Status: translation not shown
A;Molecule type: mRNA
A;Residues: 1-7 <FEE>
A;Experimental source: day 4 postnata, thymus, strain BALB/c
C;Keywords: T-cell receptor


  Query Match             33.9%;  Score 21;  DB 2;  Length 7;
  Best Local Similarity   83.3%;  Pred. No. 2.8e-05;
  Matches    5;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       6 ASSETG 11
           ||||:|
Db       1 ASSEGG 6


RESULT  13
PT0719
T-cell receptor beta chain V-D-J region (14.2.2B)  mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
C;Accession: PT0719; PT0638
R;Feeney, A.J.
J. Exp. Med. 174, 115-124, 1991
A;Title: Junctional sequences of fetal T cell receptor beta chains have few N regions
A;Reference number: PT0509; MUID:91277601
A;Accession: PT0719
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-7 <FEE>
```

A;Experimental source: newborn thymus, strain BALB/c (clone 146 2F)
A;Accession: PI0068
A;Status: translation not shown
A;Molecule type: mRNA
A;Residues: 1-?<FF2>
A;Experimental source: newborn thymus, strain BALB/c (clone 146 1N)
C;Keywords: T cell receptor

Query Match 33.9%; Score 21; DB 2; Length ?;
Best Local Similarity 84.6%; Pred. No. 2.8e+05;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy * ASSEG 11

Db ? ASSEG ?

RESULT 14
PH0818
T-cell receptor beta chain V D-J region (isolate 4) - rat (fragment)
C;Species: Rattus norvegicus (Norway rat)
C;Date: 14-Oct-1992 #sequence_revision 09-Oct-1992 #text_change 03-May-1997
C;Accession: PH0818
R;Gilfillan, E.; Otten, H.; Sun, D.; Wiley, S.; Vandenbark, A.A.; Wilson, D.B.
J. Exp. Med. 174, 1467-1476, 1991
A;Title: Analysis of T-cell receptor beta chains in Lewis rats with experimental allergi
A;Reference number: PH0891; MUID:92078857
A;Accession: PH0818
A;Molecule type: mRNA
A;Residues: 1-?<1>
A;Experimental source: concanavalin A activated lymphoblast
C;Keywords: T-cell receptor

Query Match 33.9%; Score 21; DB 2; Length ?;
Best Local Similarity 84.6%; Pred. No. 2.8e+05;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy * ASSEG 11

Db ? ASSEG ?

RESULT 15
G60787
sperm activating peptide (Ser 4,5,7 speract) - sea urchin (Anthocidaris crassispina)
C;Species: Anthocidaris crassispina
C;Date: 14-Jul-1995 #sequence_revision 04-Jun-1998 #text_change 18-Aug-2000
C;Accession: G60787
R;Suzuki, N.; Kitajima, H.; Nemura, K.; Citteri, D.L.; Yamada, K.; Kirita, M.; Tanaka, H
Comp. Biochem. Physiol. B 89, 687-693, 1988
A;Title: Some more speract derivatives associated with eggs of sea urchins, Pseudocentro
A;Reference number: A60787; MUID:88242184
A;Accession: G60787
A;Molecule type: protein
A;Residues: 1-?<SUZ>
C;Comment: This oligopeptide from egg jelly is one of several from this species, all of
a' show some but not absolute species restriction.
C;Superfamily: unassigned animal peptides

Query Match 33.9%; Score 21; DB 2; Length ?;
Best Local Similarity 84.6%; Pred. No. 2.8e+05;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy * ASSEG 11

Db ? ASSEG ?

Search completed: August 26, 2002, 06:12:50

Job time: 77 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 26, 2002, 06:28:48 ; Search time 7.5 Seconds
(without alignments)
18.930 Million cell updates/sec

Title: US-09-744-282-6
Perfect score: 62
Sequence: 1 KSFTLASSETGVG 13

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 550

Minimum DB seq length: 0
Maximum DB seq length: 13

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	20	32.3	12	1	V25K_WSSV	P82004 white spot
2	18	29.0	10	1	VEG6_BACSU	P80699 bacillus su
3	18	29.0	12	1	I69K_LOCMI	P47956 locusta mig
4	17	27.4	7	1	CIA_ENTFA	P11952 enterococcu
5	16	25.8	12	1	OPS3_DROVI	P17645 drosophila
6	15	24.2	8	1	LPK_LEIMA	P13439 leishmania
7	15	24.2	9	1	PARP_CALSI	P38495 callinectes
8	15	24.2	13	1	PSAE_PEA	P29518 pisum sativ
9	15	24.2	13	1	UN41_CLOPA	P48500 clostridium
10	14	22.6	8	1	UFO6_MOUSE	P38941 mus musculu
11	14	22.6	9	1	THYE_PIG	P11765 sus scrofa
12	14	22.6	10	1	COXO_RAT	P80452 rattus norv
13	14	22.6	10	1	KLA2_MOUSE	P39027 mus musculu
14	14	22.6	13	1	FARB_ASCSU	P45773 ascaris sum
15	13	21.0	8	1	KST_ERWCH	P50485 erwinia chr
16	13	21.0	9	1	DI_NEPNO	P24426 nephrops no
17	13	21.0	9	1	DSIP_RABIT	P01198 oryctolagus
18	13	21.0	9	1	FAR9_ASCSU	P43472 ascaris sum
19	13	21.0	10	1	CU30_LOCMI	P11785 locusta mig
20	13	21.0	10	1	IKS2_AEDAE	P42785 aedes aegyp
21	13	21.0	10	1	IK32_CHICK	P14??? gallus gal
22	13	21.0	10	1	UBA6_HUMAN	P34960 homo sapien
23	13	21.0	11	1	P90XA_CANFA	P90611 canis famil
24	13	21.0	11	1	RE41_LITRU	P82074 litoria rub
25	13	21.0	11	1	IKC2_CALVO	P41518 calliphora
26	13	21.0	12	1	CD14_LITXA	P52446 litoria xan
27	13	21.0	13	1	UN10_CLOPA	P48345 clostridium
28	12	19.4	6	1	AL14_CARMA	P81917 carcinus ma
29	12	19.4	7	1	FAR1_PROCL	P90499 procambarus
30	12	19.4	7	1	FAR2_PROCL	P90500 procambarus
31	12	19.4	8	1	AKH_LIBAU	P25418 libellula a
32	12	19.4	8	1	AKH_TABAT	P11965 tabanus atr
33	12	19.4	8	1	AL15_CARMA	P81918 carcinus ma
34	12	19.4	8	1	AL17_CARMA	P41929 carcinus ma
35	12	19.4	8	1	AL18_CARMA	P81921 carcinus ma
36	12	19.4	8	1	AL13_CYIPO	P82154 cydia pomon
37	12	19.4	8	1	AL14_CALVO	P41864 calliphora
38	12	19.4	8	1	AL14_CYIPO	P82155 cydia pomon
39	12	19.4	8	1	CAD1_ENTFA	P13208 enterococcu
40	12	19.4	8	1	FAR3_HOMAM	P11486 homarus ame
41	12	19.4	8	1	FAR4_HOMAM	P11487 homarus ame
42	12	19.4	8	1	HTF2_PEKAM	P47440 periplaneta
43	12	19.4	8	1	OKMY_OKOCT	P29455 octopoteus
44	12	19.4	9	1	CCAP_CARMA	P80560 carcinus ma
45	12	19.4	10	1	AL19_CARMA	P81922 carcinus ma

ALIGNMENTS

RESULT 1
V25K_WSSV
ID V25K_WSSV STANDARD; PRT; 12 AA.
AC P82004;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE 25 kDa structural polyprotein (Fragment).
OS White spot syndrome virus (WSSV).
OC Viruses; Unassigned viruses.
OX NCBI_TaxID=92652;
RN [1]
RP SEQUENCE.
RC STRAIN=South Carolina;
RX MEDLINE=20214217; PubMed=10752552;
RA Wang Q., Poulos B.T., Lightner D.V.;
RT "Protein analysis of geographic isolates of shrimp white spot syndrome virus.";
RL Arch. Virol. 145:263-274(2000).
CC -!- FUNCTION: STRUCTURAL COMPONENT OF THE VIRION.
FT NON_TER 12 12
SQ SEQUENCE 12 AA; 1283 MW; C5409A09ECB731A9 CRC64;

Query Match 32.3%; Score 20; DB 1; Length 12;
Best Local Similarity 80.0%; Pred. No. 5.8e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 SFTLA 6
|||:|
Db 4 SFTLS 8

RESULT 2
VEG6_BACSU
ID VEG6_BACSU STANDARD; PRT; 1 AA.
AC P80699;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Vegetative protein 6 (VEG6) (Fragment).
GN YZZI.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillus/Clostridium group
OC Bacillus/Staphylococcus group; Bacillus
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE.
RC STRAIN=IS58;
RX MEDLINE=97237728; PubMed=9084183;
RA Schmid R., Bernhardt J., Antelmann H., Voelker U., Mach H.,
RA Voelker A., Becker M.;
RT "Identification of vegetative proteins for a two dimensional protein
RT index of Bacillus subtilis.";
RL Microbiology 143:991-998(1997).
DR SubtiList; BG19024; yzzi.

```
KW   Sex pheromone; pheromone.
FT   NON_TER       1       1
SQ   SEQUENCE   12 AA;  1444 MW;  9B7C1AD8C4E6AAAA CRC64;


  Query Match             29.0%;  Score 28;  DB 1;  Length 12;
  Best Local Similarity   57.1%;  Pred. No. 1.4e+04;
  Matches    4;  Conservative    1;  Mismatches    2;  Indels    1;  Gaps    0;

Qy        1 AASEIF 6
            |:|:|
Db        1 AGSDIF 6


RESULT   4
CIA_ENCFA
ID   CIA_ENCFA      STANDARD;      PRT;     7 AA.
AC   P21695;
DT   01-OCT-1990 (Rel. 18, Created)
DT   01-OCT-1990 (Rel. 18, Last sequence update)
DT   01-FEB-1996 (Rel. 33, Last annotation update)
DE   Sex pheromone cAM373 (clumping inducing agent) (CIA).
OS   Enterococcus faecalis (Streptococcus faecalis).
OC   Bacteria; Firmicutes; Bacillus/Clostridium group; Enterococcaceae;
OC   Enterococcus.
OX   NCBI_TaxID=1351;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=87005262; PubMed=3093276;
RA   Mori M., Tanaka H., Sakagami Y., Isogai A., Fujino M., Kitada C.,
RA   Watanabe A., Ando F.Y., Clewell D.B., Suzuki A.;
RT   "Isolation and structure of the Streptococcus faecalis sex pheromone,
RT   cAM373.";
RL   FEBS Lett. 206:69-72(1986).
CC   -!- FUNCTION: CAM373 INDUCES MATING RESPONSE IN RECIPIENT CELLS
CC       HARBORING CAM373.
CC   -!- MISCELLANEOUS: THE N-TERMINUS IS PARTICULARLY RESPONSIBLE FOR
CC       SPECIFICITY OF PHEROMONES TO PLASMIDS.
CC   -!- SIMILARITY: C-TERMINAL TO THE PHEROMONES CPD1 AND CAD1.
DR   PIR; A25269; A25269.
KW   Pheromone.
SQ   SEQUENCE   7 AA;  734 MW;  75BDD72059C059D0 CRC64;


  Query Match             27.4%;  Score 17;  DB 1;  Length 7;
  Best Local Similarity   80.0%;  Pred. No. 1e+05;
  Matches    4;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy        5 FTIAS 7
            | |||
Db        5 FTIAS 7


RESULT   5
OPS4_DROVI
ID   OPS4_DROVI      STANDARD;      PRT;     12 AA.
AC   P17645;
DT   01-AUG-1990 (Rel. 15, Created)
DT   01-AUG-1990 (Rel. 15, Last sequence update)
DT   01-FEB-1996 (Rel. 33, Last annotation update)
DE   Opsin Rh4 (Inner R7 photoreceptor cells opsin) (Fragment).
GN   RH4.
OS   Drosophila virilis (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7244;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=90249748; PubMed=2140105;
RA   Fortini M.E., Rubin G.M.;
RT   "Analysis of cis-acting requirements of the Rh3 and Rh4 genes reveals
RT   a bipartite organization to rhodopsin promoters in Drosophila
RT   melanogaster.";
RL   Genes Dev. 4:444-463(1990).
CC   -!- FUNCTION: VISUAL PIGMENTS ARE GPCRS WHICH ARE RESPONSIBLE FOR
CC       MEDIATE VISION. THEY CONSIST OF AN APOPROTEIN, OPSIN, COVALENTLY
CC       LINKED TO CIS-RETINAL.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- MISCELLANEOUS: EACH DROSOPHILA EYE IS COMPOSED OF 800 FACETS OR
CC       OMMATIDIA. EACH OMMATIDIUM CONTAINS 8 PHOTORECEPTOR CELLS (R1-R8).
CC       THE R1 TO R6 CELLS ARE OUTER CELLS, WHILE R7 AND R8 ARE INNER
CC       CELLS.
CC   -!- MISCELLANEOUS: OPSIN RH4 IS SENSITIVE TO UV LIGHT.
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G PROTEIN-COUPLED RECEPTORS.
CC       OPSIN SUBFAMILY.
CC   ---------- . . . . . .
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------- . . . . . .
DR   EMBL; X51350; CAA35742.1; -.
DR   GCRDb; GCR_0779; -.
DR   FlyBase; FBgn0013091; Dvir\Rh4.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   InterPro; IPR001760; opsin.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; PARTIAL.
DR   PROSITE; PS00238; OPSIN; PARTIAL.
KW   Photoreceptor; Retinal protein; Transmembrane; Phosphorylation;
KW   Glycoprotein; G-protein coupled receptor; Vision.
FT   CARBOHYD      10      10       N-LINKED (GlcNAc...) (POTENTIAL).
FT   NON_TER       12      12
```

```
SQ   SEQUENCE   12 AA;  1253 MW;  04024E4349586B80 CRC64;


  Query Match             25.8%;  Score 16;  DB 1;  Length 12;
  Best Local Similarity   50.0%;  Pred. No. 3.4e+03;
  Matches    2;  Conservative    2;  Mismatches    0;  Indels    1;  Gaps    1;

Qy      10 IRG 13
           : :
Db       9 SSIS 5


RESULT   6
LPK_LEUMA
ID   LPK_LEUMA      STANDARD;      PRT;     9 AA.
AC   P18049;
DT   01-JAN-1993 (Rel. 25, Created)
DT   01-FEB-1994 (Rel. 28, Last sequence update)
DT   01-FEB-1994 (Rel. 28, Last annotation update)
DE   Leucopyrokinin (LPK) (LEM PE).
OS   Leucophaea maderae (Madeira cockroach).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Orthopteroidea; Dictyoptera; Blattaria;
OC   Blaberoidea; Blaberidae; Leucophaea.
OX   NCBI_TaxID=6988;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=86269041; PubMed=3015140;
RA   Nachman R.J., Holman G.M., Cook B.J.;
RT   "Active fragments and analogs of the insect neuropeptide
RT   leucopyrokinin: structure-function studies.";
RL   Biochem. Biophys. Res. Commun. 137:936-942(1986).
RN   [2]
RP   SEQUENCE, AND SYNTHESIS.
RC   TISSUE=Head;
RX   MEDLINE=87052651; PubMed=2877794;
RA   Holman G.M., Cook B.J., Nachman R.J.;
RT   "Primary structure and synthesis of a blocked myotropic
RT   neuropeptide isolated from the cockroach, Leucophaea maderae.";
RL   Comp. Biochem. Physiol. 85C:219-224(1986).
CC   -!- FUNCTION: MEDIATES VISCERAL MUSCLE CONTRACTILE ACTIVITY
CC       (MYOTROPIC ACTIVITY).
CC   -!- MISCELLANEOUS: AN ANALOG WITHOUT THE N TERMINAL PCA RESIDUE WAS
CC       SYNTHESIZED AND FOUND TO EXHIBIT GREATER ACTIVITY (11FS) THAN THE
CC       PARENT NEUROPEPTIDE. THE PORTION OF THE SEQUENCE FOR LPK MOST
CC       CRITICAL FOR THE MYOTROPIC PROPERTIES IS LIMITED TO THE
CC       PENTAPEPTIDE FRAGMENT FTPRL.
CC   -!- SIMILARITY: BELONGS TO THE PYROKININ FAMILY.
DR   PIR; A23967; A23967.
DR   InterPro; IPR001484; Pyrokinin.
DR   PROSITE; PS00539; PYROKININ; 1.
KW   Neuropeptide; Amidation; Pyrokinin.
FT   MOD_RES       1       1       PYRROLIDONE CARBOXYLIC ACID.
FT   MOD_RES       8       8       AMIDATION.
SQ   SEQUENCE   9 AA;  949 MW;  92341771A9B5A186 CRC64;


  Query Match             24.2%;  Score 15;  DB 1;  Length 8;
  Best Local Similarity   100.0%;  Pred. No. 1e+05;
  Matches    3;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       4 SPI 6
           |||
Db       5 SPI 7


RESULT   7
FARP_FALSI
ID   FARP_FALSI     STANDARD;      PRT;     9 AA.
AC   P38445;
DT   01-OCT-1994 (Rel. 30, Created)
DT   01-OCT-1994 (Rel. 30, Last sequence update)
DT   02-NOV-1995 (Rel. 32, Last annotation update)
DE   FMRFamide like neuropeptide.
OS   Callinectes sapidus (Blue crab).
OC   Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC   Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Brachyura;
OC   Eubrachyura; Portunoidea; Portunidae; Callinectes.
OX   NCBI_TaxID=6763;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=92270479; PubMed=1615219;
RA   Krajniak K.G.;
RT   "The identification and structure activity relationships of a
RT   cardioactive FMRFamide-related peptide from the blue crab Callinectes
RT   sapidus.";
RL   Peptides 12:1295-1302(1991).
CC   -!- FUNCTION: CARDIOACTIVE PEPTIDE.
CC   -!- SIMILARITY: BELONGS TO THE FARP (FMRFAMIDE RELATED PEPTIDE)
CC       FAMILY.
KW   Neuropeptide; Amidation.
FT   MOD_RES       9       9       AMIDATION.
SQ   SEQUENCE   9 AA;  1159 MW;  134F072961A4045B CRC64;


  Query Match             24.2%;  Score 15;  DB 1;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 1e+05;
  Matches    3;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 RSF 3
           |||
Db       4 RSF 6


RESULT   8
PSAE_PEA
ID   PSAE_PEA       STANDARD;      PRT;    15 AA.
AC   P20118;
DT   01-FEB-1991 (Rel. 17, Created)
DT   01-FEB-1991 (Rel. 17, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Photosystem I reaction center subunit IV (PSI-E) (Fe-S system I 15 kDa
DE   protein) (Fragment).
GN   PSAE.
OS   Pisum sativum (Garden pea).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids I; Fabales; Fabaceae; Papilionoideae; Vicieae; Pisum.
OX   NCBI_TaxID=3888;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=88137587; PubMed=3277667;
RA   Dunn P.P.J., Packman L.C., Pappin D., Gray J.C.;
RT   "N-terminal amino acid sequence analysis of the subunits of pea
RT   photosystem I.";
RL   FEBS Lett. 228:157-161(1988).
CC   -!- FUNCTION: MAY FORM COMPLEXES WITH FERREDOXIN AND FERREDOXIN
CC       OXIDOREDUCTASE IN PHOTOSYSTEM I (PS I) REACTION CENTRE.
CC   -!- SUBCELLULAR LOCATION: Chloroplast thylakoid membrane.
CC   -!- SIMILARITY: BELONGS TO THE PSAE FAMILY.
DR   PIR; S00316; S00316.
KW   Photosynthesis; Photosystem I; Chloroplast; Thylakoid; Membrane.
FT   NON_TER      13      13
SQ   SEQUENCE   13 AA;  1190 MW;  1E95A63A5C9CBAE7 CRC64;


  Query Match             24.2%;  Score 15;  DB 1;  Length 13;
  Best Local Similarity   60.0%;  Pred. No. 2e+05;
  Matches    3;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       6 ASSET 10
           | : :|
Db       1 ASEDT 5


RESULT   9
```

```
ENTRY  [illegible]
ID   [illegible]     STANDARD;      PRT;     [illegible] AA.
AC   [illegible];
DT   [illegible] (Rel. 46, Created)
DT   [illegible] (Rel. 46, Last sequence update)
DT   [illegible] (Rel. 46, Last annotation update)
DE   [illegible] protein OP 41 from 2D page (Fragment).
OS   Clostridium pasteurianum.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
OC   Clostridium.
OX   NCBI_TaxID=1953;
RN   [1]
RP   SEQUENCE.
RC   STRAIN W5;
RX   MEDLINE 96299118; PubMed 8629918;
RA   Fleissner E., Skjeldal L.;
RT   "Two-dimensional gel electrophoresis separation and N-terminal
RT   sequence analysis of proteins from Clostridium pasteurianum W5.";
RL   Electrophoresis 17:802-806(1996).
CC   -!- MISCELLANEOUS: ON THE 2D GEL THE DETERMINED PI OF THIS UNKNOWN
CC       PROTEIN IS 5.6, ITS MW IS: 29.5 kDa.
FT   NON_TER       16     16
SQ   SEQUENCE   16 AA;  1402 MW;  [illegible] CRC64;


  Query Match              24.2%;  Score 15;  DB 2;  Length 16;
  Best Local Similarity    44.4%;  Pred. No. [illegible];
  Matches    4;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 [illegible] 12

Db         [illegible]


RESULT 11
[illegible]
ID   [illegible]     STANDARD;      PRT;     8 AA.
AC   P38044;
DT   01-[illegible] (Rel. 30, Created)
DT   01-[illegible] (Rel. 30, Last sequence update)
DT   01-[illegible] (Rel. 41, Last annotation update)
DE   Unknown protein from 2D page of fibroblasts (P50) (Fragment).
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciuromorpha; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Fibroblast;
RX   MEDLINE 95000007; PubMed 7528108;
RA   Mori K.A., Patterson P.M., Wichter L.L., He Y., Solkirk L.K.;
RT   "Separation and sequencing of familiar and novel murine proteins
RT   [illegible] preparative two-dimensional gel electrophoresis.";
RL   Electrophoresis 15:735-745(1994).
CC   -!- MISCELLANEOUS: ON THE 2D GEL THE DETERMINED PI OF THIS UNKNOWN
CC       PROTEIN IS: 5.2, ITS MW IS: [illegible] kDa.
FT   NON_TER       8      8
SQ   SEQUENCE   8 AA;  817 MW;  A3DD8726760E0B1 CRC64;


  Query Match              22.6%;  Score 14;  DB 1;  Length 8;
  Best Local Similarity    75.0%;  Pred. No. [illegible];
  Matches    3;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       7 [illegible] 11
             | | |
Db       2 [illegible] 5

RESULT 12
[illegible]
ID   [illegible]     STANDARD;      PRT;     9 AA.
AC   P[illegible];
DT   21-JUL-1986 (Rel. 01, Created)
DT   21-JUL-1986 (Rel. 01, Last sequence update)
DT   21-JUL-1986 (Rel. 01, Last annotation update)
DE   Thymic factor.
OS   Sus scrofa (Pig).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX   NCBI_TaxID=9823;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE 78026571; PubMed 91486?;
RA   Pleau J. M., Dardenne M., Blouquit Y., Bach J.-F.;
RT   "Structural study of circulating thymic factor: a peptide isolated
RT   from pig serum. II. Amino acid sequence.";
RL   J. Biol. Chem. 252:8045-8047(1977).
CC   -!- MISCELLANEOUS: THE BIOLOGICAL SOURCE(S) AND PHYSIOLOGICAL
CC       ACTIVITIES OF THIS PEPTIDE HAVE NOT BEEN DETERMINED.
DR   PIR; A01523; YFPG.
FT   MOD_RES       1      1       PYRROLIDONE CARBOXYLIC ACID.
SQ   SEQUENCE   9 AA;  876 MW;  [illegible] CRC64;


  Query Match              22.6%;  Score 14;  DB 1;  Length 9;
  Best Local Similarity    60.0%;  Pred. No. [illegible];
  Matches    3;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       6 ASSEIG 11
             | | |
Db       2 AKSQGG 7


RESULT 13
COXC_RAT
ID   COXC_RAT       STANDARD;      PRT;     10 AA.
AC   P80432;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
DE   Cytochrome c oxidase polypeptide VIIc, mitochondrial (EC 1.9.3.1)
DE   (VIIIA) (Fragment).
GN   COX7C OR COX7C1.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciuromorpha; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE.
RC   STRAIN WISTAR; TISSUE=Liver, and Heart;
RX   MEDLINE 95324529; PubMed 7601105;
RA   Schaesqer H., Noack H., Hafanqk W., Brandt U., von Jagow G.;
RT   "Cytochrome-c oxidase in developing rat heart. Enzyme properties and
RT   amino-terminal sequences suggest identity of the fetal heart and the
RT   adult liver isoform.";
RL   Eur. J. Biochem. 230:235-241(1995).
CC   -!- FUNCTION: THIS PROTEIN IS ONE OF THE NUCLEAR-CODED POLYPEPTIDE
CC       CHAINS OF CYTOCHROME C OXIDASE, THE TERMINAL OXIDASE IN
CC       MITOCHONDRIAL ELECTRON TRANSPORT.
CC   -!- CATALYTIC ACTIVITY: 4 ferrocytochrome c + O(2) = 4 ferricytochrome
CC       c + 2 H(2)O.
CC   -!- SIMILARITY: BELONGS TO THE CYTOCHROME C OXIDASE VIIC FAMILY.
KW   Oxidoreductase; Mitochondrion.
FT   NON_TER       10     10
SQ   SEQUENCE   10 AA;  1117 MW;  12CDE767687B1B0B CRC64;


  Query Match              22.6%;  Score 14;  DB 1;  Length 10;
  Best Local Similarity    60.0%;  Pred. No. 6.7e-04;
  Matches    3;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       9 EIGVG 13
             | | |
Db       4 EEGPG 8
```

```
RESULT  13
RLA2_MOUSE
ID   RLA2_MOUSE     STANDARD;      PRT:    [illegible] AA.
AC   [illegible];
DT   15-OCT-1998 (Rel. 37, Created)
DT   15-OCT-1998 (Rel. 37, Last sequence update)
DT   15-OCT-1998 (Rel. 37, Last annotation update)
DE   60S acidic ribosomal protein P2 (Fragment).
GN   RPLP2.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID [illegible];
RN   [1]
RP   SEQUENCE.
RC   TISSUE [illegible];
RA   Sanchez J., [illegible], Rouge V., Frutinger S., Hughes G., Yan J.X.,
RA   Hoogland C., Appel R.D., Binz P.-A., Hochstrasser D.F.,
RA   Cawthorne M.;
RL   Submitted (AUG 1998) to the SWISS-PROT data bank.
CC   -!- FUNCTION: PLAYS AN IMPORTANT ROLE IN THE FI[illegible] S[illegible] P
CC       PROTEIN SYNTHESIS.
CC   -!- SUBUNIT: P1 AND P2 EXIST AS DIMERS AT THE LARGE RIB[illegible] MAL
CC       SUBUNIT.
CC   -!- PTM: PHOSPHORYLATED (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE L12P FAMILY OF RIBOSOMAL PROTEINS.
DR   SWISS-2DPAGE; P90027; MOUSE.
KW   Ribosomal protein; Phosphorylation.
FT   NON_TER       10      10
SQ   SEQUENCE   10 AA;  1186 MW;  07121E3B45BDC2DB CRC64;


  Query Match             22.6%;  Score 14;  DB 1;  Length 10;
  Best Local Similarity   60.0%;  Pred. No. 6.7e+04;
  Matches    3;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       2 SFILA 6
           |:  |
Db       6 SYLLA 10


RESULT  14
FARP_ASCSU
ID   FARP_ASCSU     STANDARD;      PRT:    13 AA.
AC   P34[illegible];
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   [illegible]-FEB-1999 (Rel. 33, Last annotation update)
DE   FMRFamide like neuropeptide AF11.
OS   Ascaris suum (Pig roundworm) (Ascaris lumbricoides).
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Ascaridida; Ascaridoidea;
OC   Ascarididae; Ascaris.
OX   NCBI_TaxID [illegible];
RN   [1]
RP   SEQUENCE.
RX   MEDLINE [illegible]; PubMed 7651904;
RA   Cowden C., Stretton A.O.W.;
RT   "Eight novel FMRFamide-like neuropeptides isolated from the nematode
RT   Ascaris suum.";
RL   Peptides 16:341-[illegible] (1995).
CC   -!- SIMILARITY: BELONGS TO THE FARP (FMRFAMIDE RELATED PEPTIDE)
CC       FAMILY.
KW   Neuropeptide; Amidation.
FT   MOD_RES       13      13       AMIDATION.
SQ   SEQUENCE   13 AA;  1195 MW;  9CAEC650D6886B05 CRC64;


  Query Match             22.6%;  Score 14;  DB 1;  Length 13;
  Best Local Similarity   40.0%;  Pred. No. 8.8e+04;
  Matches    2;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       8 SFIV 11
           :  |
Db       [illegible]; [illegible];
Db       1 SDIGI 5


RESULT  15
RS1_ERWCH
ID   RS1_ERWCH      STANDARD;      PRT:    8 AA.
AC   P37985;
DT   01-OCT-1994 (Rel. 30, Created)
DT   01-OCT-1994 (Rel. 30, Last sequence update)
DT   01-NOV-1995 (Rel. 32, Last annotation update)
DE   30S ribosomal protein S1 (Fragment).
GN   RPSA.
OS   Erwinia chrysanthemi.
OC   Bacteria; Proteobacteria; gamma subdivision; Enter[illegible]bacteriaceae;
OC   Pectobacterium.
OX   NCBI_TaxID 556;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=3937;
RA   Douillie A., Toussaint A., Faelon M.;
RL   Submitted (AUG 1994) to the EMBL/GenBank/DDBJ databases.
CC   !  FUNCTION: BINDS MRNA, THUS FACILITATING RECOGNITION OF THE
CC      INITIATION POINT. IT IS NEEDED TO TRANSLATE MRNA WITH A SHORT
CC      SHINE DALGARNO (SD) PURINE RICH SEQUENCE (BY SIMILARITY).
CC   !  SIMILARITY: BELONGS TO THE S1 FAMILY OF RIBOSOMAL PROTEINS.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See [illegible] for
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; X74750; CAA52769.1; -.
DR   PIR; S37141; S37141.
KW   Ribosomal protein; Repeat; RNA-binding.
FT   NON_TER       1       1
SQ   SEQUENCE   8 AA;  837 MW;  [illegible]DC9B33C1 CRC64;


  Query Match             21.0%;  Score 13;  DB 1;  Length 8;
  Best Local Similarity   37.5%;  Pred. No. [illegible];
  Matches    3;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       2 SFILASSE 9
           :|  |  |
Db       1 AFKSAKGE 8


Search completed: August 26, 2002, [illegible]
Job time: 601 sec
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

SW protein : protein search, using sw model

Run on: August 26, 2002, 06:23:01 ; Search time 81.17 Seconds
(without alignments)
27.436 Million cell updates/sec

Title: US-09-744-282-6
Perfect score: 21
Sequence: 1 RSPIIASSEIGVG 13

Scoring table: BL_SUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 2252

Minimum DB seq length: 0
Maximum DB seq length: 13

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_19:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	21	33.9	9	15	Q85710	Q85710 ROUS SARCOM
2	21	33.9	13	13	Q98922	Q98922 BRASSICA OL
3	19	30.6	12	2	Q50442	Q50442 MYCOPLASMA
4	18	29.0	11	2	Q94M17	Q94M17 ESCHERICHIA
5	18	29.0	12	11	Q9WUX1	Q9WUX1 RATTUS NORV
6	18	29.0	12	13	Q9PS67	Q9PS67 CAVIA PALI
7	18	29.0	13	14	P82880	P82880 RANA CHLAMI
8	17	27.4	8	2	Q9X6D5	Q9X6D5 ESCHERICHIA
9	17	27.4	8	3	P82858	P82858 PACHYSTIMA RE
10	17	27.4	9	7	Q9MW43	Q9MW43 HOMO SAPIEN
11	17	27.4	11	11	Q90XN6	Q90XN6 MUS MUSCULU
12	17	27.4	11	11	Q99JC3	Q99JC3 RATTUS SP.
13	17	27.4	13	4	O75905	O75905 HOMO SAPIEN
14	17	27.4	13	4	Q9HAQ9	Q9HAQ9 HOMO SAPIEN
15	16	25.8	8	2	P72221	P72221 PSEUDOMONAS
16	16	25.8	8	4	Q9P285	Q9P285 HOMO SAPIEN
17	16	25.8	8	11	Q90615	Q90615 MUS MUSCULU
18	16	25.8	9	2	Q93567	Q93567 SALMONELLA
19	16	25.8	9	6	Q9XS10	Q9XS10 CAPRA HIRC
20	16	25.8	9	11	Q9QVH9	Q9QVH9 MUS SP. SUP
21	16	25.8	9	11	Q9QZA8	Q9QZA8 MUS MUSCUL
22	16	25.8	10	6	Q95M70	Q95M70 TRICHOSURUS
23	16	25.8	12	6	Q91RV6	Q91RV6 BOS TAURUS
24	16	25.8	12	16	Q41858	Q41858 ZEA MAYS (M
25	16	25.8	13	2	Q9R514	Q9R514 STREPTOCOCC
26	16	25.8	13	4	Q9UDE6	Q9UDE6 HOMO SAPIEN
27	16	25.8	13	4	Q9UET1	Q9UET1 HOMO SAPIEN
28	16	25.8	13	15	P82064	P82064 TINEOCYCLUS
29	16	25.8	13	15	Q98YM1	Q98YM1 HUMAN HUMOR
30	15	24.2	7	2	P70804	P70804 AZOTOBACTER
31	15	24.2	8	5	Q946Z3	Q946Z3 MARIPOSA VIR
32	15	24.2	8	8	P92422	P92422 PSATHYROSTA
33	15	24.2	8	8	P92878	P92878 BAYNEUDINA C
34	15	24.2	8	8	P93985	P93985 AEQUIPUS CO
35	15	24.2	8	8	P92404	P92404 LOPHOPYRUM
36	15	24.2	8	8	P92429	P92429 PSEUDOTSUGA
37	15	24.2	8	8	P93979	P93979 CHAMOPYRUM
38	15	24.2	8	8	P93973	P93973 CHAMOPYRUM
39	15	24.2	8	8	P92388	P92388 BOTRARDIA P
40	15	24.2	8	8	P92428	P92428 PERIDITYUS
41	15	24.2	8	8	P92401	P92401 HETERANTHET
42	15	24.2	8	8	P92227	P92227 CRITHOPSIS
43	15	24.2	8	8	P93993	P93993 PSATHYROSTA
44	15	24.2	8	8	P93961	P93961 PSATHYROSTA
45	15	24.2	8	8	P92219	P92219 ANTIOPYRUM

ALIGNMENTS

RESULT 1
Q85710
ID Q85710 PRELIMINARY; PRT; 9 AA.
AC Q85710;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE POL PROTEIN (FRAGMENT).
OS Rous sarcoma virus.
OC Viruses; Retroid viruses; Retroviridae; Alpharetrovirus.
OX NCBI_TaxID=11886;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=84115080; PubMed=6319754;
RA Lerner T.L., Hanafusa H.;
RT "DNA sequence of the Bryan high titer strain of Rous sarcoma virus:
RT Extent of env deletion and possible genetical relationship with
RT other viral strains.";
RL J. Virol. 49:549-556(1984).
DR EMBL; K03365; AAA42557.1; -.
FT NON_TER 1 1
SQ SEQUENCE 9 AA; 949 MW; 94AA144DDDE7C1AA CRC64;

Query Match 33.9%; Score 21; DB 15; Length 9;
Best Local Similarity 57.1%; Pred. No. 5.6e+04;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 4 IIASSEIGV 12
|||:::|:
Db 5 IIAANESGV 13

RESULT 2
Q9S922
ID Q9S922 PRELIMINARY; PRT; 13 AA.
AC Q9S922;
DT 01-MAY-2000 (TrEMBLrel. 14, Created)
DT 01-MAY-2000 (TrEMBLrel. 14, Last sequence update

```
ID   [illegible] MAY 2000 (TrEMBLrel. 14, Last annotation update)
DE   [illegible] SPECIFIC GLYCOPROTEIN.
OS   Brassica oleracea (Cauliflower).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Brassicales; Brassicaceae; Brassica.
OX   NCBI_TaxID=[illegible];
RN   [1]
RP   SEQUENCE.
RX   MEDLINE 9[illegible]; PubMed=175224[illegible];
RA   Grado L., Demoroy L., Dumas C.;
RT   "Use of a fast protein electrophoretic purification procedure for N
RT   terminal sequence analysis to identify S locus related proteins in
RT   stigmas of Brassica oleracea.";
RL   Electrophoresis 12:646-654(1991).
SQ   SEQUENCE   [illegible] AA;  1451 MW;  96C24EDFF09060B5 CRC64;


  Query Match             33.3%;  Score 21;  DB 10;  Length 14;
  Best Local Similarity   57.1%;  Pred. No. 2.6e+04;
  Matches    4;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 LAGSET 6
           : ::|:|
Db       6 LSSSIPS 12


RESULT   5
Q[illegible]42
ID   Q[illegible]42       PRELIMINARY;       PRT;    12 AA.
AC   Q[illegible]42;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   ATP SYNTHASE ALPHA (FRAGMENT).
OS   Mycoplasma pneumoniae.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
OC   Mycoplasmataceae; Mycoplasma.
OX   NCBI_TaxID=2104;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=M129;
RX   MEDLINE=[illegible]; PubMed=7984113;
RA   Proft T., Herrmann R.;
RT   "Identification and characterization of hitherto unknown Mycoplasma
RT   pneumoniae proteins.";
RL   Mol. Microbiol. 13:337-348(1994).
DR   EMBL; [illegible]; CAA83503.1; -.
FT   NON_TER       12      12
SQ   SEQUENCE   12 AA;  1496 MW;  8A7DE4E61CA1A681 CRC64;


  Query Match             30.6%;  Score 19;  DB 2;  Length 12;
  Best Local Similarity   40.0%;  Pred. No. [illegible];
  Matches    4;  Conservative    2;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       [illegible]SSET 10
           :|:
Db       [illegible]ELIMISKI 11


RESULT   4
[illegible]
ID   [illegible]       PRELIMINARY;       PRT;    11 AA.
AC   [illegible];
DT   [illegible] (TrEMBLrel. 10, Created)
DT   [illegible] (TrEMBLrel. 19, Last sequence update)
DT   [illegible] (TrEMBLrel. 19, Last annotation update)
DE   ADP-RIBOSYLATION FACTOR (FRAGMENT).
GN   [illegible].
OS   [illegible].
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciuromorpha; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=98345789; PubMed=9672068;
RA   Greeve J., Axelos D., Welker S., Schipper M., Greten H.;
RT   "Distinct promoters induce APOBEC-1 expression in rat liver and
RT   intestine.";
RL   Arterioscler. Thromb. Vasc. Biol. 18:1079-1092(1998).
DR   EMBL; AJ006695; CAB44489.1; -.
FT   NON_TER       1       1
FF   NON_TER       12      12
SQ   SEQUENCE   12 AA;  1357 MW;  70FB167969932588 CRC64;


  Query Match             29.0%;  Score 18;  DB 11;  Length 12;
  Best Local Similarity   100.0%;  Pred. No. 7.9e+03;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       7 SSET 10
           ||||
Db       9 SSET 12


RESULT   6
Q9PS67
ID   Q9PS67       PRELIMINARY;       PRT;    12 AA.
AC   Q9PS67;
DT   01-MAY-2000 (TrEMBLrel. 14, Created)
DT   01-MAY-2000 (TrEMBLrel. 14, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   MYOSIN.
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=9031;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=92034444; PubMed=1931944;
RA   Garbodian T.E., Yount R.G.;
RT   "Direct photoaffinity labeling of gizzard myosin with vanadate trapped
```

```
RT   adenosine diphosphate.";
RL   Biochemistry 30:10126-10132(1991).
SQ   SEQUENCE   12 AA;  1178 MW;  C1FC2B0D9AA3487D CRC64;


  Query Match             29.0%;  Score 18;  DB 13;  Length 12;
  Best Local Similarity   60.0%;  Pred. No. 7.9e+03;
  Matches    7;  Conservative    1;  Mismatches    2;  Indels    1;  Gaps    0;

Qy       9 EIGVG 13
           |:|:|
Db       2 ESGAG 6


RESULT   7
P82880
ID   P82880      PRELIMINARY;      PRT;    13 AA.
AC   P82880;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   TEMPORIN-1CA.
OS   Rana clamitans (green frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Neobatrachia; Ranoidea; Ranidae; Rana.
OX   NCBI_TaxID=145282;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=SKIN;
RX   MEDLINE=20284865; PubMed=10822131;
RA   Halverson T., Basir Y.J., Knoop F.C., Conlon J.M.;
RT   "Purification and characterization of antimicrobial peptides from the
RT   skin of the North American green frog Rana clamitans.";
RL   Peptides 21:469-476(2000).
CC   -!- FUNCTION: ANTIBACTERIAL ACTIVITY AGAINST GRAM-POSITIVE BACTERIUM
CC       S.AUREUS.
CC   -!- SUBCELLULAR LOCATION: SECRETED.
CC   -!- MASS SPECTROMETRY: MW=1420.0; MW_ERR=0.02; METHOD=ELECTROSPRAY.
CC   -!- SIMILARITY: BELONGS TO THE BREVININ/TEMPORIN/ESCULENTIN/BRACOVIN
CC       FAMILY.
KW   Antibiotic; Amidation.
FT   MOD_RES      13     13       AMIDATION.
SQ   SEQUENCE   13 AA;  1492 MW;  D2195329B6FB2783 CRC64;


  Query Match             29.0%;  Score 18;  DB 13;  Length 13;
  Best Local Similarity   62.5%;  Pred. No. 8.6e+03;
  Matches    5;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       5 LASSEIGV 12
           |:| :||
Db       5 LAKILIGV 12


RESULT   8
Q9S6D5
ID   Q9S6D5      PRELIMINARY;      PRT;     8 AA.
AC   Q9S6D5;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   PUTATIVE IS30 TRANSPOSASE (FRAGMENT).
OS   Escherichia coli.
OC   Bacteria; Proteobacteria; Gamma subdivision; Enterobacteriaceae;
OC   Escherichia.
OX   NCBI_TaxID=562;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=A295;
RX   MEDLINE=99194747; PubMed=10094716;
RA   Kuhn A., Drummelsmith J., Whitfield C.;
RT   "Conserved organization in the cps gene clusters for expression of
RT   Escherichia coli serogroup K5 antigens: relationship to the colanic acid
RT   biosynthesis locus and the cps genes from Klebsiella pneumoniae.";
RL   J. Bacteriol. 181:2307-2314(1999).
DR   EMBL; AF118251; AAD30008.1; -.
FT   NON_TER      8      8
SQ   SEQUENCE   8 AA;  1011 MW;  F21DC1A9D1B4414F CRC64;


  Query Match             27.4%;  Score 17;  DB 2;  Length 8;
  Best Local Similarity   75.0%;  Pred. No. 1.6e+04;
  Matches    3;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 RSPT 4
           |:||
Db       3 RIPT 6


RESULT   9
P82858
ID   P82858      PRELIMINARY;      PRT;     8 AA.
AC   P82858;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   URICASE (EC 1.7.3.3) (URATE OXIDASE) (FRAGMENT).
OS   Puccinia recondita f. sp. triseti.
OC   Eukaryota; Fungi; Basidiomycota; Urediniomycetes; Urediniomycetidae;
OC   Uredinales; Pucciniaceae; Puccinia.
OX   NCBI_TaxID=142679;
RN   [1]
RP   SEQUENCE, FUNCTION, AND DEVELOPMENTAL STAGE.
RC   TISSUE=SPORE;
RA   Aguilar M., Montalbini P., Pineda M.;
RL   Submitted (NOV-2000) to the EMBL/GenBank/DDBJ databases.
CC   !  FUNCTION: INVOLVED IN HOST-PARASITE RELATIONSHIP BETWEEN PLANTS
CC      AND FUNGI.
CC   -!- CATALYTIC ACTIVITY: URATE + O(2) + H(2)O = 5-HYDROXYISOURATE +
CC       H(2)O(2) (5-HYDROXYISOURATE DECOMPOSES TO FORM ALLANTOIN).
CC   -!- SUBCELLULAR LOCATION: PEROXISOMAL (BY SIMILARITY).
CC   -!- DEVELOPMENTAL STAGE: GERMINATION.
CC   -!- SIMILARITY: BELONGS TO THE URICASE FAMILY.
DR   InterPro; IPR002042; Uricase.
DR   PROSITE; PS00366; URICASE; PARTIAL.
KW   Oxidoreductase; Purine metabolism; Peroxisome.
FT   NON_TER      8      8
SQ   SEQUENCE   8 AA;  777 MW;  98C1ADD735B9D76D CRC64;


  Query Match             27.4%;  Score 17;  DB 2;  Length 8;
  Best Local Similarity   50.0%;  Pred. No. 5.6e+05;
  Matches    3;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3 FTLASS 8
           |:|::
Db       3 FSLATA 8


RESULT  10
Q9MW43
ID   Q9MW43      PRELIMINARY;      PRT;     9 AA.
AC   Q9MW43;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   MHC CLASS I ANTIGEN (FRAGMENT).
GN   HLA-B39061.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=BLOOD;
RX   MEDLINE=97246738; PubMed=9089104;
```

```
RA   Varela Alarcon C., Gomez Casado E., Martinez Laso J., Granados J.,
RA   Larrad Sainz Z., Alegre R., Arnaiz Villena A.;
RT   "Differences in intron 2 sequences between B*3906 and B*3905 in
RT   Amerindians: comparison with those of B*3901, B*5101, and B*5201 2
RT   alleles.";
RL   Immunogenetics 45:436-439(1997).
DR   EMBL; L76640; AAF86298.1; -.
FT   NON_TER       1       1
FT   NON_TER       9       9
SQ   SEQUENCE   9 AA;  975 MW;  51C661AB01A72D07 CRC64;


  Query Match             27.4%;  Score 17;  DB 7;  Length 9;
  Best Local Similarity   66.7%;  Pred. No. 5.2e+05;
  Matches    4;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       5 LASSEI 10
           : :||
Db       2 LALIEI 7


RESULT  12
Q9QXN6
ID   Q9QXN6      PRELIMINARY;      PRT;    11 AA.
AC   Q9QXN6;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   INTERLEUKIN 2 (FRAGMENT).
GN   IL2.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=2;
RX   MEDLINE=20248845; PubMed=10779485;
RA   Lyons P.A., Armitage N., Argentina F., Denny P., Hill N.J., Lord C.J.,
RA   Wilson M.B., Peterson L.B., Wicker L.S., Todd J.A.;
RT   "Congenic mapping of the type 1 diabetes locus, idd3, to a 780-kb
RT   region of mouse chromosome 3: identification of a candidate segment of
RT   ancestral DNA by haplotype mapping.";
RL   Genome Res. 10:446-453(2000).
DR   EMBL; AF105695; AAF22748.1; -.
DR   MGD; MGI:96548; Il2.
FT   NON_TER      11      11
SQ   SEQUENCE   11 AA;  1234 MW;  AE24E9C9BDB736D6 CRC64;


  Query Match             27.4%;  Score 17;  DB 11;  Length 11;
  Best Local Similarity   55.6%;  Pred. No. 1.1e+04;
  Matches    5;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       2 SEILASSEI 10
           :  |:: :|
Db       3 SMQLASCVI 11


RESULT  13
Q9CC73
ID   Q9CC73      PRELIMINARY;      PRT;    11 AA.
AC   Q9CC73;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   DELETING HOMEOBOX-CONTAINING GENE 2 PROTEIN (FRAGMENT).
DE   (FRAGMENT).
OS   Rattus sp.
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10118;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=OVARY;
RX   MEDLINE=96147985; PubMed=8571710;
RA   Shen Q.X., Liu H.H., Chen W.Y., Bahl O.P.;
RT   "[Cloning and overexpression of rat ovary LH/hCG receptor cDNA in
RT   insect cells].";
RL   Shih Yen Sheng Wu Hsueh Pao 28:283-290(1995).
DR   EMBL; S80658; AAB50709.1; -.
KW   Receptor; Chorion.
FT   NON_TER       1       1
FT   NON_TER      11      11
SQ   SEQUENCE   11 AA;  994 MW;  333DCB137EB865B8 CRC64;


  Query Match             27.4%;  Score 17;  DB 11;  Length 11;
  Best Local Similarity   60.0%;  Pred. No. 1.1e+04;
  Matches    3;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       9 ETGVG 13
           :|| |
Db       2 QTGSG 6


RESULT  13
O75905
ID   O75905      PRELIMINARY;      PRT;    13 AA.
AC   O75905;
DT   01-NOV-1998 (TrEMBLrel. 08, Created)
DT   01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   PROSAPOSIN (FRAGMENT).
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=98426155; PubMed=9751800;
RA   Sun Y., Jin P., Witte D.P., Grabowski G.A.;
RT   "Isolation and characterization of the human prosaposin promoter.";
RL   Gene 218:37-47(1998).
DR   EMBL; AF057307; AAC64921.1; -.
FT   NON_TER      13      13
SQ   SEQUENCE   13 AA;  1413 MW;  4B285B6C17841727 CRC64;


  Query Match             27.4%;  Score 17;  DB 4;  Length 13;
  Best Local Similarity   80.0%;  Pred. No. 1.3e+04;
  Matches    4;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       3 FTLAS 7
           | |||
Db       5 FLLAS 9


RESULT  14
Q9HAQ9
ID   Q9HAQ9      PRELIMINARY;      PRT;    13 AA.
AC   Q9HAQ9;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   PROSAPOSIN (FRAGMENT).
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21205764; PubMed=11309366;
RA   Bukovska G., Kostolanska K., Jezova M., Jezova J., Ivanchenko M., Bircak A.,
RA   Poupetova H., Befekadu A., Berna L., Paton B.C., Harzer K., Bouvet A.,
RA   Smid F., Elleder M.;
```

```
RT   "A 1.2-kb deletion in the coding region of the prosaposin gene leads to
RT   a severe deficiency of prosaposin and saposins, and is associated
RT   with a complex sphingolipidosis dominated by lactosylceramide
RT   accumulation.";
RL   Hum. Mol. Genet. 10:927-940(2001).
DR   EMBL; AF327850; AAG36645.1; -.
FT   NON_TER       13      13
SQ   SEQUENCE   13 AA;  1384 MW;  57985B6C178407E7 CRC64;


  Query Match             27.4%;  Score 17;  DB 4;  Length 13;
  Best Local Similarity   80.0%;  Pred. No. 1.3e+04;
  Matches    4;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       5 FILAS 7
           |:|:|
Db       5 FLLAS 9


RESULT  15
P72221
ID   P72221      PRELIMINARY;      PRT;     8 AA.
AC   P72221;
DT   01-FEB-1997 (TrEMBLrel. 02, Created)
DT   01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT   01-OCT-2001 (TrEMBLrel. 19, Last annotation update)
DE   ALGINATE LYASE (FRAGMENT).
GN   ALY.
OS   Pseudomonas sp. (strain OS ALY-1).
OC   Bacteria; Proteobacteria.
OX   NCBI_TaxID=86668;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=OS ALY-1;
RA   Fujiyama K.;
RL   Submitted (OCT-1994) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=OS-ALY-1;
RX   MEDLINE=96329966; PubMed=8836113;
RA   Maki H., Mori A., Fujiyama K., Kinoshita S., Yoshida T.;
RT   "Cloning, sequence analysis and expression in Escherichia coli of a
RT   gene coding for an alginate lyase from Pseudomonas sp. OS ALY-1.";
RL   J. Gen. Microbiol. 139:987-993(1993).
DR   EMBL; D38466; BAA21704.1;
KW   Lyase.
FT   NON_TER       1       1
FT   NON_TER       8       8
SQ   SEQUENCE   8 AA;  841 MW;  461DDDC5A5804180 CRC64;


  Query Match             25.8%;  Score 16;  DB 2;  Length 8;
  Best Local Similarity   88.9%;  Pred. No. 5.6e+05;
  Matches    2;  Conservative    4;  Mismatches    0;  Indels    1;  Gaps    0;

Qy       4 FILASS 9
           :|||:|
Db       1 YLLSSA 6



Search completed: August 26, 2002, 06:23:02
Job time: 609 sec
```

GeneCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

-M protein - protein search, using sw model

Run on: August 26, 2002, 06:16:29 ; Search time 80.0 Seconds
(without alignments)
17.895 Million cell updates/sec

Title: US-09-744-282-6
Perfect score: 62
Sequence: 1 KSFILASSEIVG 13

Scoring table: blosum62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 100000

Minimum DB seq length: 0
Maximum DB seq length: 13

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_042802:*
1: /cgn2_6/ptodata/1/pubpep/pep_nucl_AA:*
2: /cgn2_6/ptodata/1/pubpep/pep_nucl_AB:*
3: /cgn2_6/ptodata/1/pubpep/pep_nucl_AC:*
4: /cgn2_6/ptodata/1/pubpep/pep_nucl_AD:*
5: /cgn2_6/ptodata/1/pubpep/pep_nucl_AE:*
6: /cgn2_6/ptodata/1/pubpep/pep_nucl_AF:*
7: /cgn2_6/ptodata/1/pubpep/pep_nucl_AG:*
8: /cgn2_6/ptodata/1/pubpep/pep_nucl_AH:*
9: /cgn2_6/ptodata/1/pubpep/pep_nucl_AI:*
10: /cgn2_6/ptodata/1/pubpep/pep_nucl_AJ:*
11: /cgn2_6/ptodata/1/pubpep/pep_nucl_AK:*
12: /cgn2_6/ptodata/1/pubpep/pep_nucl_AL:*
13: /cgn2_6/ptodata/1/pubpep/pep_nucl_AM:*
14: /cgn2_6/ptodata/1/pubpep/pep_nucl_AN:*
15: /cgn2_6/ptodata/1/pubpep/pep_nucl_AO:*
16: /cgn2_6/ptodata/1/pubpep/pep_nucl_AP:*
17: /cgn2_6/ptodata/1/pubpep/pep_nucl_AQ:*
18: /cgn2_6/ptodata/1/pubpep/pep_nucl_AR:*
19: /cgn2_6/ptodata/1/pubpep/pep_nucl_AS:*
20: /cgn2_6/ptodata/1/pubpep/pep_nucl_AT:*
21: /cgn2_6/ptodata/1/pubpep/pep_nucl_AU:*
22: /cgn2_6/ptodata/1/pubpep/pep_nucl_AV:*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	62	100.0	13	17	AAR95076	Peptide contg. hC
2	62	100.0	13	19	AAW61597	Human cartilage op
3	62	100.0	13	21	AAY69231	Peptide fragment c
4	62	100.0	13	21	AAY77401	Human cartilage ol
5	62	100.0	13	22	AAU03045	Human cartilage op
6	62	100.0	13	22	AAU03048	Human cartilage op
7	62	100.0	13	22	AAU03059	Human cartilage op
8	62	100.0	13	22	AAU03065	Human cartilage op
9	62	100.0	13	22	AAU03066	Human cartilage op
10	62	100.0	13	22	AAU03067	Human cartilage op
11	58	93.5	13	22	AAU03054	Human cartilage op
12	58	93.5	13	22	AAU03073	Human cartilage op
13	57	91.9	13	22	AAU03046	Human cartilage op
14	57	91.9	13	22	AAU03047	Human cartilage op
15	57	91.9	13	22	AAU03052	Human cartilage op
16	57	91.9	13	22	AAU03057	Human cartilage op
17	57	91.9	13	22	AAU03058	Human cartilage op
18	57	91.9	13	22	AAU03062	Human cartilage op
19	57	91.9	13	22	AAU03064	Human cartilage op
20	57	91.9	13	22	AAU03072	Human cartilage op
21	55	88.7	13	22	AAU03050	Human cartilage op
22	55	88.7	13	22	AAU03051	Human cartilage op
23	55	88.7	13	22	AAU03056	Human cartilage op
24	55	88.7	13	22	AAU03060	Human cartilage op
25	55	88.7	13	22	AAU03061	Human cartilage op
26	55	88.7	13	22	AAU03063	Human cartilage op
27	53	85.5	13	22	AAU03053	Human cartilage op
28	47	75.8	13	22	AAU03049	Human cartilage op
29	46	74.2	13	21	AAY77400	Human chondrocyte
30	46	74.2	13	22	AAU03055	Human cartilage op
31	43	69.4	9	17	AAR95072	Human cartilage ol
32	43	69.4	9	21	AAY69227	Peptide fragment c
33	42	67.7	13	21	AAB03459	Spork derived coil
34	41	66.1	13	17	AAR95075	Peptide contg. hC
35	41	66.1	13	21	AAY69230	Peptide fragment c
36	34	54.8	9	21	AAY77399	Human chondrocyte
37	32	51.6	9	17	AAR95071	Human cartilage ol
38	32	51.6	9	21	AAY69226	Peptide fragment c
39	27	43.5	13	21	AAB26177	Trichoderma reesei
40	26	41.9	11	22	AAM98235	Human peptide #33
41	26	41.9	13	21	AAY66638	E coli antigen rec
42	24	38.7	6	6	AAB71240	Synthetic apoptii
43	24	38.7	12	14	AAW13113	Primer# 1 relate
44	24	38.7	12	14	AAB29070	Aphanocladium albu
45	24	38.7	12	18	AAW40635	Peptide which bind

ALIGNMENTS

RESULT 1
AAR95076
ID AAR95076 standard; peptide; 13 AA.
XX
AC AAR95076;
XX
DT 17-JAN-1996 (first entry)
XX
DE Peptide contg. hC up-39 antagonist used in pept. Bo 111.
XX
KW Acute cartilage glycoprotein 39; hC up-39; autoimmune;
KW infertility; T cell tolerance mediated destruction;
KW rheumatoid arthritis; disease model; fibrosis.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Peptide 3..11
FT /note "hC up-39 peptide (11)"
XX
PN WO9613517-A1.
XX
PD 09-MAY-1996.
XX
PF 25-OCT-1995; 95WO-EP04201.
XX
PR 07-APR-1995; 95EP-0200886.
PR 27-OCT-1994; 94EP-0203128.
XX
PA (ALKU) AKZO NOBEL NV.
XX
PI Boots AMH, Verheijden GFM;
XX
DR WPI; 1996-239447/24.

```
XX
PT   Peptide(s) derived from human cartilage gp 39 and fragments    used
PT   to induce specific T cell tolerance, esp. for treatment of
PT   rheumatoid arthritis, also useful in producing arthritic animal
PT   model
XX
PS   Claim 27; Page 48; 47pp; English.
XX
CC   The present peptide, contg. a peptide derived from the human
CC   cartilage glycoprotein 39 (HC gp 39) autoantigen, can be used to
CC   induce specific T cell tolerance to HC gp 39 in a patient with
CC   T cell mediated cartilage destruction, esp. rheumatoid arthritis
CC   (RA). It may also be used to induce RA in mice, etc. (human disease
CC   model) and to diagnose the presence of activated T cells
CC   that react with HC gp 39, i.e. to identify patients suitable for
CC   T cell tolerant treatment.
CC   Therapeutically, the peptide is given in a dose of 0.1 to microg/kg
CC   by injection, or 10-50 microg/kg to produce a disease model. The
CC   peptide has a specific tolerising effect on autoreactive T cells,
CC   but leaves other components of the immune system intact, and
CC   is said to have low side effects.
XX
SQ   Sequence    13 AA;


  Query Match             100.0%;  Score 62;  DB 19;  Length 13;
  Best Local Similarity   100.0%;  Pred. No. 4.2e-05;
  Matches   13;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 RSFILASSETGVG 13
           |||||||||||||
Db       1 rsfilassetqvq 13


RESULT    3
AAY69231
ID   AAY69231 standard; peptide; 13 AA.
XX
AC   AAY69231;
XX
DT   30-MAY-2000  (first entry)
XX
DE   Peptide fragment of human cartilage gp 39 (HC gp 39) protein.
XX
KW   Human cartilage gp 39 protein; HC gp 39 protein; T cell activator;
KW   inflammatory disease; lymphocyte; antigen; bystander suppression;
KW   antigen-specific T-cell tolerance; autoimmune disorder;
KW   inflammatory disease; rheumatoid arthritis;
KW   autoimmune inflammatory disease; autoantigen.
XX
OS   Homo sapiens.
XX
PN   WO200004917-A2.
XX
PD   03-FEB-2000.
XX
PF   19-JUL-1999;   99WO-EP05331.
XX
PR   23-JUL-1998;   98EP-0202471.
XX
PA   (ALEX ) ARGEN-NOBEL NV.
XX
PI   Mittenburg AMM,  Boots AMH;
XX
DR   WPI; 2000-182587/16.
XX
PT   Use of cartilage protein or peptides for treating inflammatory diseases
PT   especially rheumatoid arthritis
XX
PS   Claim 2; Page 17; 29pp; English.
XX
CC   AAY69226-33 represent fragments of the human cartilage gp 39 (HC gp 39)
CC   protein. The peptides are T-cell activators. The peptides are used for
CC   preventing inflammatory diseases, by modulating lymphocyte reactive
CC   to antigens other than HC gp 39, present in the same tissue of
CC   HC gp 39. By the induction of antigen specific T cell tolerance,
CC   autoimmune disorders can be treated by bystander suppression. The
CC   HC gp 39 peptides are useful for preventing inflammatory diseases
CC   especially rheumatoid arthritis. The fragments may also be used for
CC   treating autoimmune inflammatory diseases. The HC gp 39 peptides are
CC   especially beneficial in autoimmune conditions in which the
CC   autoantigens associated with the condition are unknown.
XX
SQ   Sequence    13 AA;
```

Query Match 100.0%; Score 62; DB 21; Length 13;
Best Local Similarity 100.0%; Pred. No. 4.2e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RSFTLASSETGVG 13
|||||||||||||
Db 1 rsftlassetqvq 13

RESULT 4
AAY77401
ID AAY77401 standard; peptide; 13 AA.
XX
AC AAY77401;
XX
DT 22-MAY-2000 (first entry)
XX
DE Human cartilage glycoprotein-39 (HC gp-39), residues 265-275.
XX
KW Human cartilage glycoprotein 39; HC gp-39; YKL-40; epitope;
KW autoreactive T-cell; tolerisation; desensitisation; autoantigen;
KW rheumatoid arthritis; human chondrocyte protein; YKL-39;
KW bystander suppression; autoimmune disorder; Graves' disease;
KW primary glomerulonephritis; inflammatory bowel disease; diagnostic.
XX
OS Homo sapiens.
XX
PN WO200005254-A2.
XX
PD 03-FEB-2000.
XX
PF 16-JUL-1999; 99WO-EP05050.
XX
PR 23-JUL-1998; 98EP-0202470.
XX
PA (AKZO) AKZO NOBEL NV.
XX
PI Verhelst GFM, Boots AMH;
XX
DR WPI; 2000-184744/16.
XX
PT New T-cell epitope peptides useful for inducing specific T-cell
PT tolerance to an autoantigen for treating autoimmune disorders,
PT especially rheumatoid arthritis -
XX
PS Example 1; Page 10; 26pp; English.
XX
CC The invention relates to peptides constituting amino acids 268-276 and
CC amino acids 266-278 of human chondrocyte protein YKL-39
CC (AAY77399-Y77400, respectively). The peptides of the invention may be
CC used in pharmaceutical compositions for the induction of specific T-cell
CC tolerance (desensitisation) to an autoantigen in patients
CC suffering from autoimmune disorders, especially rheumatoid arthritis. A
CC frequently found autoantigen in rheumatoid arthritis sufferers is human
CC cartilage glycoprotein-39 (HC gp-39, also known as YKL-40). Several
CC epitopes of HC gp-39 are recognised by peripheral blood T cells from
CC rheumatoid arthritis patients; one such epitope, HC gp-39 residues
CC 265-275, was recognised by T cells from 8 out of 18 patients tested.
CC YKL-39 shares significant sequence identity to HC gp-39. YKL-39 residues
CC 266-278 (AAY77400) constitutes a mimicry epitope of HC gp-39 (265-275),
CC and may be used for desensitisation of T-cells with reactivity to HC
CC gp-39 (265-275) in rheumatoid arthritis patients. Rheumatoid arthritis
CC may be treated by the systemic administration of compositions comprising
CC the peptides of the invention, optionally in combination with existing
CC treatments. The peptides may also be used for modulating lymphocytes
CC that are reactive to autoantigens other than YKL-39 and HC gp-39 but
CC which are present in the same tissue as these autoantigens. By the
CC induction of antigen-specific T-cell tolerance, autoimmune disorders can
CC be treated by bystander suppression. The conditions that may be treated
CC via this approach include Graves' disease, primary glomerulonephritis
CC and inflammatory bowel disease. The peptides, alone with a detection
CC agent, may additionally be used in a diagnostic composition. The peptides
CC of the invention have a specific effect on autoreactive T-cells whilst
CC leaving other components of the immune system intact. Therefore, the
CC peptides should be safe, having no side effects. The present sequence
CC represents residues 265-275 of human cartilage glycoprotein 39 (HC gp-39)
CC which was used in an exemplification of the present invention.
XX
SQ Sequence 13 AA;

Query Match 100.0%; Score 62; DB 21; Length 13;
Best Local Similarity 100.0%; Pred. No. 4.2e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RSFTLASSETGVG 13
|||||||||||||
Db 1 rsftlassetqvq 13

RESULT 5
AAU03045
ID AAU03045 standard; peptide; 13 AA.
XX
AC AAU03045;
XX
DT 12-SEP-2001 (first entry)
XX
DE Human cartilage gp-39 autoimmune peptide #1.
XX
KW Human; autoimmune disease; epitope; antigen; rheumatoid arthritis;
KW immunosuppression.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Modified-site 1
FT /note= "Deaminated"
XX
PN WO200129081-A1.
XX
PD 26-APR-2001.
XX
PF 12-OCT-2000; 2000WO-EP10230.
XX
PR 18-OCT-1999; 99EP-0204427.
XX
PA (AKZO) AKZO NOBEL NV.
XX
PI Van Staveren CJ, Timmers CM, Van Galen PJM, Kroeze EMA, Boots AMH;
PI Miltenburg AMM;
XX
DR WPI; 2001-367293/38.
XX
PT Modified peptides derived from human cartilage gp-39 protein epitope
PT are useful in immunotherapy, especially in the treatment of arthritis
PT -
XX
PS Claim 1; Page 36; 52pp; English.
XX
CC The sequence represents Human cartilage gp-39 protein epitope (amino
CC acids 265-277) which is used as the template for deamidated modified
CC peptides for use in immunotherapy. The peptides are useful for inducing
CC tolerance induction in patients suffering from autoimmune diseases such
CC as rheumatoid arthritis. The peptides provide an effective, non-toxic
CC immunosuppression therapy as a very attractive alternative to
CC non-specific immunosuppression.
XX
SQ Sequence 13 AA;

Query Match 100.0%; Score 62; DB 22; Length 13;
Best Local Similarity 100.0%; Pred. No. 4.2e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RSFTLASSETGVG 13

```
          ; sttllllall
Db        1 rsttlassetqvq 13


RESULT   7
AA003064
ID   AA003064 standard; peptide; 13 AA.
XX
AC   AA003064;
XX
DT   12-SEP-2001  (first entry)
XX
DE   Human cartilage gp-39 autoimmune peptide #4.
XX
KW   Human; autoimmune disease; epitope; antigen; rheumatoid arthritis;
KW   immunosuppression.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT                   /note= "N-terminal (N-methyl-nicotinoyl)-"
FT   Modified-site   13
FT                   /note= "C-terminus is amidated"
XX
PN   WO200129081-A1.
XX
PD   26-APR-2001.
XX
PF   12-OCT-2000; 2000WO-EP10230.
XX
PR   18-OCT-1999;  99EP-0203427.
XX
PA   (ASTE-) ASTRAZENECA NV.
XX
PI   Van Staveren CJ, Timmers CM, Van Galen PJM, Knoester RMA, Boots AMH;
PI   Miltenburg AMM;
XX
DR   WPI; 2001-367293/38.
XX
PT   Modified peptides derived from human cartilage gp-39 protein epitope
PT   are useful in immunotherapy, especially in the treatment of arthritis
PT
XX
PS   Claim 12; Page 39; 52pp; English.
XX
CC   The sequence represents a modified peptide derived from Human cartilage
CC   gp-39 protein epitope (amino acids 263-275) for use in immunotherapy.
CC   The peptides are useful for inducing tolerance induction in patients
CC   suffering from autoimmune diseases such as rheumatoid arthritis.
CC   The peptides provide antigen specific, non toxic immunosuppression
CC   therapy as a very attractive alternative to non specific
CC   immunosuppression.
XX
SQ   Sequence   13 AA;


  Query Match             100.0%;  Score 62;  DB 22;  Length 13;
  Best Local Similarity   100.0%;  Pred. No. 4.2e-05;
  Matches   13;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 RSFILASSETGVG 13
           |||||||||||||
Db       1 rsttlassetqvq 13


RESULT   8
AA003065
ID   AA003065 standard; peptide; 13 AA.
XX
AC   AA003065;
XX
DT   12-SEP-2001  (first entry)
XX
DE   Human cartilage gp-39 autoimmune peptide #21.
XX
KW   Human; autoimmune disease; epitope; antigen; rheumatoid arthritis;
KW   immunosuppression.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
```

FT /note "N terminal acetyl"
FT Misc-difference 1
FT /note "D-from residue"
XX
PN WO200129081-A1.
XX
PD 26-APR-2001.
XX
PF 12-OCT-2000; 2000WO-EP10230.
XX
PR 18-OCT-1999; 99EP-0203427.
XX
PA (ALKU) AKZO NOBEL NV.
XX
PI Van Staveren CJ, Timmers CM, Van Galen PJM, Knegt-L RMA, Boots AMH;
PI Miltenburg AMM;
XX
DR WPI; 2001-367293/38.
XX
PT Modified peptides derived from human cartilage gp-39 protein epitope
PT are useful in immunotherapy, especially in the treatment of arthritis
PT -
XX
PS Disclosure; Page 34; 52pp; English.
XX
CC The sequence represents a modified peptide derived from Human cartilage
CC gp-39 protein epitope (amino acids 263-275) for use in immunotherapy.
CC The peptides are useful for inducing tolerance induction in patients
CC suffering from autoimmune diseases such as rheumatoid arthritis.
CC The peptides provide antigen-specific, non-toxic immunosuppression
CC therapy as a very attractive alternative to non-specific
CC immunosuppression.
XX
SQ Sequence 13 AA;

Query Match 100.0%; Score 62; DB 22; Length 13;
Best Local Similarity 100.0%; Pred. No. 4.2e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RSFTLASSETGVG 13
|||||||||||||
Db 1 rsftlassetgvg 13

RESULT 9
AAU03066
ID AAU03066 standard; peptide; 13 AA.
XX
AC AAU03066;
XX
DT 12-SEP-2001 (first entry)
XX
DE Human cartilage gp-39 autoimmune peptide #22.
XX
KW Human; autoimmune disease; epitope; antigen; rheumatoid arthritis;
KW immunosuppression.
XX
OS Homo sapiens.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Modified-site 1
FT /note "N terminal acetyl"
XX
PN WO200129081-A1.
XX
PD 26-APR-2001.
XX
PF 12-OCT-2000; 2000WO-EP10230.
XX
PR 18-OCT-1999; 99EP-0203427.
XX
PA (ALKU) AKZO NOBEL NV.
XX
PI Van Staveren CJ, Timmers CM, Van Galen PJM, Knegt-L RMA, Boots AMH;
PI Miltenburg AMM;
XX
DR WPI; 2001-367293/38.
XX
PT Modified peptides derived from human cartilage gp-39 protein epitope
PT are useful in immunotherapy, especially in the treatment of arthritis
PT -
XX
PS Example 15; Page 32; 52pp; English.
XX
CC The sequence represents a modified peptide derived from Human cartilage
CC gp-39 protein epitope (amino acids 263-275) for use in immunotherapy.
CC The peptides are useful for inducing tolerance induction in patients
CC suffering from autoimmune diseases such as rheumatoid arthritis.
CC The peptides provide antigen-specific, non-toxic immunosuppression
CC therapy as a very attractive alternative to non-specific
CC immunosuppression.
XX
SQ Sequence 13 AA;

Query Match 100.0%; Score 62; DB 22; Length 13;
Best Local Similarity 100.0%; Pred. No. 4.2e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RSFTLASSETGVG 13
|||||||||||||
Db 1 rsftlassetgvg 13

RESULT 10
AAU03067
ID AAU03067 standard; peptide; 13 AA.
XX
AC AAU03067;
XX
DT 12-SEP-2001 (first entry)
XX
DE Human cartilage gp-39 autoimmune peptide #23.
XX
KW Human; autoimmune disease; epitope; antigen; rheumatoid arthritis;
KW immunosuppression.
XX
OS Homo sapiens.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Modified-site 1
FT /note "N terminal Met-Tyr(O)"
XX
PN WO200129081-A1.
XX
PD 26-APR-2001.
XX
PF 12-OCT-2000; 2000WO-EP10230.
XX
PR 18-OCT-1999; 99EP-0203427.
XX
PA (ALKU) AKZO NOBEL NV.
XX
PI Van Staveren CJ, Timmers CM, Van Galen PJM, Knegt-L RMA, Boots AMH;
PI Miltenburg AMM;
XX
DR WPI; 2001-367293/38.
XX
PT Modified peptides derived from human cartilage gp-39 protein epitope
PT are useful in immunotherapy, especially in the treatment of arthritis
PT -
XX
PS Example 9; Page 25; 52pp; English.

```
XX
CC   The sequence represents a modified peptide derived from Human cartilage
CC   gp-39 protein epitope (amino acids 263-275) for use in immunotherapy.
CC   The peptides are useful for inducing tolerance induction in patients
CC   suffering from autoimmune diseases such as rheumatoid arthritis.
CC   The peptides provide antigen specific, non toxic immunosuppression
CC   therapy as a very attractive alternative to non specific
CC   immunosuppression..
XX
SQ   Sequence   13 AA;


  Query Match             100.0%;  Score 62;  DB 22;  Length 13;
  Best Local Similarity   100.0%;  Pred. No. 4.2e-05;
  Matches   13;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 RSFTLASSETGVG 13
           |||||||||||||
Db       1 rsftlassetgvg 13


RESULT  11
AAU00054
ID   AAU00054 standard; peptide; 13 AA.
XX
AC   AAU00054;
XX
DT   12-SEP-2001  (first entry)
XX
DE   Human cartilage gp-39 autoimmune peptide #10.
XX
KW   Human; autoimmune disease; epitope; antigen; rheumatoid arthritis;
KW   immunosuppression.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT                   /note= "N-terminal acetyl"
FT   Modified-site   2
FT                   /label= OTHER
FT                   /note= "Other= NhSer"
FT   Modified-site   13
FT                   /note= "C-terminus is amidated"
XX
PN   WO200129081-A1.
XX
PD   26-APR-2001.
XX
PF   12-OCT-2000; 2000WO-EP10230.
XX
PR   18-OCT-1999;   99EP-0203427.
XX
PA   (ALKU ) AKZO NOBEL NV.
XX
PI   Van Staveren CJ,  Timmers CM,  Van Galen PJM,  Knegtel RMA,  Boots AMH;
PI   Miltenburg AMM;
XX
DR   WPI; 2001-367293/38.
XX
PT   Modified peptides derived from human cartilage gp-39 protein epitope
PT   are useful in immunotherapy, especially in the treatment of arthritis
PT
XX
PS   Claim 12; Page 99; 52pp; English.
XX
CC   The sequence represents a modified peptide derived from Human cartilage
CC   gp-39 protein epitope (amino acids 263-275) for use in immunotherapy.
CC   The peptides are useful for inducing tolerance induction in patients
CC   suffering from autoimmune diseases such as rheumatoid arthritis.
CC   The peptides provide antigen specific, non toxic immunosuppression
CC   therapy as a very attractive alternative to non specific
CC   immunosuppression.
XX
SQ   Sequence   13 AA;


  Query Match             93.5%;  Score 58;  DB 22;  Length 13;
  Best Local Similarity   92.3%;  Pred. No. 0.00024;
  Matches   12;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 RSFTLASSETGVG 13
           | ||||||||||||
Db       1 rxftlassetqvq 13


RESULT  12
AAU03073
ID   AAU03073 standard; peptide; 13 AA.
XX
AC   AAU03073;
XX
DT   12-SEP-2001  (first entry)
XX
DE   Human cartilage gp-39 autoimmune peptide #29.
XX
KW   Human; autoimmune disease; epitope; antigen; rheumatoid arthritis;
KW   immunosuppression.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT                   /note= "N-terminal acetyl"
FT   Modified-site   2
FT                   /label= OTHER
FT                   /note= "Other= NhSer"
FT   Modified-site   13
FT                   /note= "C-terminal is amidated"
XX
PN   WO200129081-A1.
XX
PD   26-APR-2001.
XX
PF   12-OCT-2000; 2000WO-EP10230.
XX
PR   18-OCT-1999;   99EP-0203427.
XX
PA   (ALKU ) AKZO NOBEL NV.
XX
PI   Van Staveren CJ,  Timmers CM,  Van Galen PJM,  Knegtel RMA,  Boots AMH;
PI   Miltenburg AMM;
XX
DR   WPI; 2001-367293/38.
XX
PT   Modified peptides derived from human cartilage gp-39 protein epitope
PT   are useful in immunotherapy, especially in the treatment of arthritis
PT
XX
PS   Example 15; Page 32; 52pp; English.
XX
CC   The sequence represents a modified peptide derived from Human cartilage
CC   gp-39 protein epitope (amino acids 263-275) which is used as the template
CC   for use in immunotherapy. The peptides are useful for inducing
CC   tolerance induction in patients suffering from autoimmune diseases such
CC   as rheumatoid arthritis. The peptides provide antigen specific, non toxic
CC   immunosuppression therapy as a very attractive alternative to
CC   non-specific immunosuppression.
XX
SQ   Sequence   13 AA;


  Query Match             93.5%;  Score 58;  DB 22;  Length 13;
  Best Local Similarity   92.3%;  Pred. No. 0.00024;
```

```
  Matches   12;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 RSFTLASSETGVG 13
           | ||||.|||||||
Db       1 rxftlassetqvq 13


RESULT  13
AAU03046
ID   AAU03046 standard; peptide; 13 AA.
XX
AC   AAU03046;
XX
DT   12-SEP-2001  (first entry)
XX
DE   Human cartilage gp-39 autoimmune peptide #2.
XX
KW   Human; autoimmune disease; epitope; antigen; rheumatoid arthritis;
KW   immunosuppression.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT                   /label= OTHER
FT                   /note= "Other  desaminoargininyl"
FT   Modified-site   13
FT                   /note= "C-terminus is amidated"
XX
PN   WO200129081-A1.
XX
PD   26-APR-2001.
XX
PF   12-OCT-2000; 2000WO-EP10230.
XX
PR   18-OCT-1999;   99EP-0203427.
XX
PA   (ALKU ) AKZO NOBEL NV.
XX
PI   Van Staveren CJ,  Timmers CM,  Van Galen PJM,  Knottel RMA,  Boots AMH;
PI   Miltenburg AMM;
XX
DR   WPI; 2001-367293/38.
XX
PT   Modified peptides derived from human cartilage gp-39 protein epitope
PT   are useful in immunotherapy, especially in the treatment of arthritis
PT   -
XX
PS   Claim 12; Page 39; 52pp; English.
XX
CC   The sequence represents a modified peptide derived from Human cartilage
CC   gp-39 protein epitope (amino acids 263-275) for use in immunotherapy.
CC   The peptides are useful for inducing tolerance induction in patients
CC   suffering from autoimmune diseases such as rheumatoid arthritis.
CC   The peptides provide antigen-specific, non-toxic immunosuppression
CC   therapy as a very attractive alternative to non-specific
CC   immunosuppression.
XX
SQ   Sequence   13 AA;


  Query Match             91.9%;  Score 57;  DB 22;  Length 13;
  Best Local Similarity   100.0%;  Pred. No. 0.00037;
  Matches   12;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2 SFTLASSETGVG 13
           ||||||||||||
Db       2 sftlassetqvq 13


RESULT  14
AAU03047
ID   AAU03047 standard; peptide; 13 AA.
XX
AC   AAU03047;
XX
DT   12-SEP-2001  (first entry)
XX
DE   Human cartilage gp-39 autoimmune peptide #3.
XX
KW   Human; autoimmune disease; epitope; antigen; rheumatoid arthritis;
KW   immunosuppression.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT                   /label= OTHER
FT                   /note= "Other= desaminoargininyl"
XX
PN   WO200129081-A1.
XX
PD   26-APR-2001.
XX
PF   12-OCT-2000; 2000WO-EP10230.
XX
PR   18-OCT-1999;   99EP-0203427.
XX
PA   (ALKU ) AKZO NOBEL NV.
XX
PI   Van Staveren CJ,  Timmers CM,  Van Galen PJM,  Knottel RMA,  Boots AMH;
PI   Miltenburg AMM;
XX
DR   WPI; 2001-367293/38.
XX
PT   Modified peptides derived from human cartilage gp-39 protein epitope
PT   are useful in immunotherapy, especially in the treatment of arthritis
PT   -
XX
PS   Claim 12; Page 39; 52pp; English.
XX
CC   The sequence represents a modified peptide derived from Human cartilage
CC   gp-39 protein epitope (amino acids 263-275) for use in immunotherapy.
CC   The peptides are useful for inducing tolerance induction in patients
CC   suffering from autoimmune diseases such as rheumatoid arthritis.
CC   The peptides provide antigen-specific, non-toxic immunosuppression
CC   therapy as a very attractive alternative to non-specific
CC   immunosuppression..
XX
SQ   Sequence   13 AA;


  Query Match             91.9%;  Score 57;  DB 22;  Length 13;
  Best Local Similarity   100.0%;  Pred. No. 0.00037;
  Matches   12;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2 SFTLASSETGVG 13
           ||||||||||||
Db       2 sftlassetqvq 13


RESULT  15
AAU03052
ID   AAU03052 standard; peptide; 13 AA.
XX
AC   AAU03052;
XX
DT   12-SEP-2001  (first entry)
XX
DE   Human cartilage gp-39 autoimmune peptide #8.
XX
KW   Human; autoimmune disease; epitope; antigen; rheumatoid arthritis;
KW   immunosuppression.
XX
```

```
OS   Homo sapiens.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified site   1
FT                   /note  "N terminal acetyl"
FT   Modified site   6
FT                   /label  =OTHER
FT                   /note  "Other  NH CH(CH2CH(CH3)2) CH2"
FT   Modified site   14
FT                   /note  "C terminus is amidated"
XX
PN   WO 2002058731-A1.
XX
PD   29 AUG 2002.
XX
PF   12 JAN 2001; 2001WO EP000230.
XX
PR   18 JAN 2000;   2000EP 00300427.
XX
PA   (ARGE ) ARGENTA N BEL NV.
XX
PI   Van Eyden F;  Timmers CM;  Van Galen PJM;  Kroeze RMA;  Boots AMH;
PI   Mulder A MM;
XX
DR   WPI; 2002-743738/80.
XX
PT   Modified peptides derived from human cartilage gp 39 protein epitope
PT   are useful in immunotherapy, especially in the treatment of arthritis
PT
XX
PS   Claim 12; Page 89; 92pp; English.
XX
CC   The sequence represents a modified peptide derived from human cartilage
CC   gp 39 protein epitope (amino acids 263 275) for use in immunotherapy.
CC   The peptides are useful for inducing tolerance induction in patients
CC   suffering from autoimmune diseases such as rheumatoid arthritis.
CC   The peptides provide antigen specific, non toxic immunosuppression.
CC   therapy is a very attractive alternative to non-specific
CC   immunosuppression.
XX
SQ   Sequence   14 AA;


  Query Match             91.9%;  Score 57;  DB 22;  Length 14;
  Best Local Similarity   92.3%;  Pred. No. 0.00087;
  Matches   12;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 PSFFLASSFIGVAG 14
             |||||||||||||
Db       1 PSFFLASSFIGVAG 14


Search completed: August 26, 2002, 06:16:29
Job time : 2.0 Secs
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 26, 2002, 06:13:41 ; Search time 30.25 Seconds
(without alignments)
10,554 Million cell updates/sec

Title: US-09-744-282-6
Perfect score: 62
Sequence: 1 RSFTLASSEIGVG 13

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0
Maximum DB seq length: 13

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iraa/PCT_NEW_PUB.pep:*
2: /cgn2_6/ptodata/1/iraa/PCTUS_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/iraa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/iraa/US07_NEW_PUB.pep:*
5: /cgn2_6/ptodata/1/iraa/PCTUS_PUBCOMB.pep:*
6: /cgn2_6/ptodata/1/iraa/US06_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	62	100.0	13	1	US-08-619-645-6	Sequence 6, Appli
2	62	100.0	13	2	US-08-634-493-6	Sequence 6, Appli
3	44	71.0	9	1	US-08-619-645-2	Sequence 2, Appli
4	44	71.0	9	2	US-08-634-493-2	Sequence 2, Appli
5	41	66.1	13	1	US-08-619-645-5	Sequence 5, Appli
6	41	66.1	13	2	US-08-634-493-5	Sequence 5, Appli
7	32	51.6	9	1	US-08-619-645-1	Sequence 1, Appli
8	32	51.6	9	2	US-08-634-493-1	Sequence 1, Appli
9	24	38.7	12	1	US-07-939-501A-5	Sequence 5, Appli
10	24	38.7	12	1	US-08-250-789A-13	Sequence 13, Appl
11	24	38.7	12	1	US-08-250-789A-45	Sequence 45, Appl
12	24	38.7	12	4	US-08-448-898-9	Sequence 9, Appli
13	23	37.1	8	1	US-08-315-461-2	Sequence 2, Appli
14	23	37.1	9	4	US-09-187-859-3857	Sequence 3857, Ap
15	23	37.1	11	3	US-09-140-149-13	Sequence 13, Appl
16	23	37.1	13	4	US-09-256-050-14	Sequence 14, Appl
17	22	35.5	10	1	US-07-899-535A-2	Sequence 2, Appli
18	22	35.5	11	1	US-08-068-947-27	Sequence 27, Appl
19	22	35.5	11	4	US-09-143-708A-6	Sequence 6, Appli
20	22	35.5	12	1	US-08-250-789A-99	Sequence 99, Appl
21	22	35.5	13	4	US-08-942-456-4	Sequence 4, Appli
22	21	33.9	8	1	US-08-446-206B-19	Sequence 19, Appl
23	21	33.9	8	2	US-08-807-030-17	Sequence 17, Appl
24	21	33.9	8	4	US-09-187-859-3854	Sequence 3854, Ap
25	21	33.9	9	1	US-08-594-447-49	Sequence 49, Appl
26	21	33.9	9	1	US-08-541-964-48	Sequence 48, Appl
27	21	33.9	9	2	US-08-665-647-63	Sequence 63, Appl
28	21	33.9	9	5	PCT-US92-08054-5	Sequence 5, Appli
29	21	33.9	10	1	US-08-335-298-5	Sequence 5, Appli
30	21	33.9	10	4	US-09-187-859-3858	Sequence 3858, Ap
31	21	33.9	11	1	US-08-512-662-1	Sequence 1, Appli
32	21	33.9	11	4	US-09-187-859-3859	Sequence 3859, Ap
33	21	33.9	11	5	PCT-US96-12445-1	Sequence 1, Appli
34	21	33.9	12	1	US-08-250-789A-29	Sequence 29, Appl
35	21	33.9	12	1	US-08-408-994A-44	Sequence 44, Appl
36	21	33.9	13	1	US-08-211-979A-8	Sequence 8, Appli
37	21	33.9	13	2	US-08-807-030-7	Sequence 17, Appl
38	21	33.9	13	5	PCT-US93-10302-4	Sequence 4, Appli
39	20	32.3	5	4	US-09-187-859-156	Sequence 156, Ap
40	20	32.3	5	4	US-09-187-859-1279	Sequence 1279, Ap
41	20	32.3	6	4	US-09-187-859-159	Sequence 159, Ap
42	20	32.3	6	4	US-09-187-859-1281	Sequence 1281, Ap
43	20	32.3	7	1	US-08-136-743B-53	Sequence 53, Appl
44	20	32.3	7	3	US-09-040-219-26	Sequence 26, Appl
45	20	32.3	7	4	US-09-187-859-162	Sequence 162, App

ALIGNMENTS

RESULT 1
US-08-619-645-6
; Sequence 6, Application US/08619645
; Patent No. 5746507
; GENERAL INFORMATION:
; APPLICANT: Boots, Anna M.H.
; APPLICANT: Verheijden, Gijsbertus J.M.
; TITLE OF INVENTION: Novel peptides derived from
; TITLE OF INVENTION: autoantibodies for use in immunotherapy of rheumatoid
; FILE OF INVENTION: disease
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee Patent Department
; STREET: 1300 Piccard Drive, Suite 206
; CITY: Rockville
; STATE: Maryland
; COUNTRY: U.S.A.
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25(EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/619,645
; FILING DATE: 25-MAR-1996
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Gormley, Mary E.
; REGISTRATION NUMBER: 34,489
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 947-4433
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 13 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-619-645-6

Query Match 100.0%; Score 62; DB 2; Length 13;
Best Local Similarity 100.0%; Pred. No. 7.4e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RSFTLASSEIGVG 13
||||||||||||| |
Db 1 RSFTLASSEIGVG 13

```
RESULT   2
US-08-634-493-2
; Sequence 2, Application US/08634493
; Patent No. 5843449
;  GENERAL INFORMATION:
;    APPLICANT:  A.M.H. Boots
;    APPLICANT:  G.F.M. Verheijden
;    APPLICANT:  E.S. Bos
;    TITLE OF INVENTION:  No. 5843449cl Peptides derived from autoantigen for use
;    TITLE OF INVENTION:  in Immunotherapy of Autoimmune Diseases
;    NUMBER OF SEQUENCES:  10
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Akzo No. 5843449cl Patent Department
;      STREET:  1300 Piccard Drive, Suite 206
;      CITY:  Rockville
;      STATE:  Maryland
;      COUNTRY:  U.S.A.
;      ZIP:  20850
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25(EPO)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/619,645
;      FILING DATE:  25-MAR-1996
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Gormley, Mary E.
;      REGISTRATION NUMBER:  34,409
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (301) 947-4433
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  9 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-619-645-2


  Query Match             69.4%;  Score 43;  DB 1;  Length 9;
  Best Local Similarity   100.0%;  Prod. No. 1.7e-05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3 FILASSEIG 11
           |||||||||
Db       1 FILASSEIG 9


RESULT   4
US-08-634-493-2
; Sequence 2, Application US/08634493
; Patent No. 5843449
;  GENERAL INFORMATION:
;    APPLICANT:  A.M.H. Boots
;    APPLICANT:  G.F.M. Verheijden
;    APPLICANT:  E.S. Bos
;    TITLE OF INVENTION:  No. 5843449cl Peptides derived from autoantigen for use
;    TITLE OF INVENTION:  in Immunotherapy of Autoimmune Diseases
;    NUMBER OF SEQUENCES:  10
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Akzo No. 5843449cl Patent Department
;      STREET:  1300 Piccard Drive, Suite 206
;      CITY:  Rockville
;      STATE:  Maryland
;      COUNTRY:  USA
;      ZIP:  20850
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/634,493
;      FILING DATE:  18-APR-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/619,645
;      FILING DATE:  25-MAR-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/EP95/04201
;      FILING DATE:  25-OCT-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  NL 9420l287
```

```
;       FILING DATE:  27-OCT-1994
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  NL 95200860
;       FILING DATE:  07-APR-1995
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (301) 258-5200
;       TELEFAX:  (301) 977-0847
;   INFORMATION FOR SEQ ID NO:  2:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  9 amino acids
;       TYPE:  amino acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  peptide
US-08-644-493-2


  Query Match             69.4%;  Score 43;  DB 2;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 1.7e-05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        6 FILASSEIG 14
            |||||||||
Db        1 FILASSEIG 9


RESULT   5
US-08-619-645-5
; Sequence 5, Application US/08619645
; Patent No. 5736507
;  GENERAL INFORMATION:
;     APPLICANT:  Boots, Anna M.H.
;     APPLICANT:  Verheijden, Gijsbertus F.M.
;     TITLE OF INVENTION:  No. 5736507el peptides derived from
;     TITLE OF INVENTION:  autoantigen for use in immunotherapy of autoimmune
;     TITLE OF INVENTION:  disease
;     NUMBER OF SEQUENCES:  9
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Akzo No. 5736507el Patent Department
;       STREET:  1300 Piccard Drive, Suite 206
;       CITY:  Rockville
;       STATE:  Maryland
;       COUNTRY:  U.S.A.
;       ZIP:  20850
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25(EPO)
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/619,645
;       FILING DATE:  25-MAR-1996
;       CLASSIFICATION:  514
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Gormley, Mary E.
;       REGISTRATION NUMBER:  34,409
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (301) 947-4433
;   INFORMATION FOR SEQ ID NO:  5:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  14 amino acids
;       TYPE:  amino acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  peptide
US-08-619-645-5


  Query Match             66.1%;  Score 41;  DB 1;  Length 14;
  Best Local Similarity   100.0%;  Pred. No. 0.079;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        1 RSFILASSE 9
            |||||||||
Db        5 RSFILASSE 13


RESULT   6
US-08-644-493-5
; Sequence 5, Application US/08634493
; Patent No. 5843449
;  GENERAL INFORMATION:
;     APPLICANT:  A.M.H. Boots
;     APPLICANT:  G.F.M. Verheijden
;     APPLICANT:  E.S. Bos
;     TITLE OF INVENTION:  No. 5843449el Peptides derived from autoantigen for use
;     TITLE OF INVENTION:  in Immunotherapy of Autoimmune Diseases
;     NUMBER OF SEQUENCES:  10
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Akzo No. 5843449el Patent Department
;       STREET:  1300 Piccard Drive, Suite 206
;       CITY:  Rockville
;       STATE:  Maryland
;       COUNTRY:  USA
;       ZIP:  20850
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/634,493
;       FILING DATE:  18-APR-1996
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/619,645
;       FILING DATE:  25-MAR-1996
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  PCT/EP95/04201
;       FILING DATE:  25-OCT-1995
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  NL 942041287
;       FILING DATE:  27-OCT-1994
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  NL 95200860
;       FILING DATE:  07-APR-1995
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (301) 258-5200
;       TELEFAX:  (301) 977-0847
;   INFORMATION FOR SEQ ID NO:  5:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  14 amino acids
;       TYPE:  amino acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  peptide
US-08-644-493-5


  Query Match             66.1%;  Score 41;  DB 2;  Length 14;
  Best Local Similarity   100.0%;  Pred. No. 0.079;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        1 RSFILASSE 9
            |||||||||
Db        5 RSFILASSE 13


RESULT   7
US-08-619-645-1
; Sequence 1, Application US/08619645
; Patent No. 5736507
;  GENERAL INFORMATION:
;     APPLICANT:  Boots, Anna M.H.
;     APPLICANT:  Verheijden, Gijsbertus F.M.
;     TITLE OF INVENTION:  No. 5736507el peptides derived from
;     TITLE OF INVENTION:  autoantigen for use in immunotherapy of autoimmune
```

; TITLE OF INVENTION: disease
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Akzo No. 5730.53/01 Patent Department
; STREET: 1300 Piccard Drive, Suite 205
; CITY: Rockville
; STATE: Maryland
; COUNTRY: U.S.A.
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC DOS/MS DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25(EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/619,645
; FILING DATE: 25-MAR-1996
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Tansley, Mary E.
; REGISTRATION NUMBER: 34,400
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 947-4433
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-634-493-1

Query Match 51.6%; Score 32; DB 2; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.7e-05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 RSPILAS 7
|||||||
Db 3 RSPILAS 9

RESULT 8
US-08-634-493-1
; Sequence 1, Application US/08634493
; Patent No. 5844449
; GENERAL INFORMATION:
; APPLICANT: A.M.H. Boots
; APPLICANT: S.J.M. Verheijden
; APPLICANT: E.J. Bos
; TITLE OF INVENTION: No. 5844449el Peptides derived from autoantigen for use
; TITLE OF INVENTION: in Immunotherapy of Autoimmune Diseases
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Akzo No. 5844449el Patent Department
; STREET: 1300 Piccard Drive, Suite 206
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC DOS/MS DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/634,493
; FILING DATE: 18-APR-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/619,645
; FILING DATE: 25-MAR-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: EP/EP95/04201
; FILING DATE: 25-OCT-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: NL 9420341287
; FILING DATE: 27-OCT-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: NL 952008860
; FILING DATE: 07-APR-1995
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 258-5200
; TELEFAX: (301) 977-0847
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-634-493-1

Query Match 51.6%; Score 32; DB 2; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.7e-05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 RSPILAS 7
|||||||
Db 3 RSPILAS 9

RESULT 9
US-07-939-501A-5
; Sequence 5, Application US/07939501A
; Patent No. 5446138
; GENERAL INFORMATION:
; APPLICANT: BLAISEAU, Pierre-Louis
; APPLICANT: LEGOUX, Richard
; APPLICANT: LEGUAY, Jean-Jacques
; APPLICANT: SCHNEIDER, Michel
; TITLE OF INVENTION: Recombinant DNA coding for a protein
; TITLE OF INVENTION: having endochitinase activity
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FOLEY & LARDNER
; STREET: King Street Station,Suite 500,1800 Diagonal
; STREET: Road,P.O. Box 299
; CITY: ALEXANDRIA
; STATE: VIRGINIA
; COUNTRY: USA
; ZIP: 22314-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/939,501A
; FILING DATE: 19920908
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 91 11072
; FILING DATE: 06-SEP-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: SAXE, Bernhard D
; REGISTRATION NUMBER: 28,665
; REFERENCE/DOCKET NUMBER: 16281/473
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-9300
; TELEFAX: (703) 683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 12 amino acids
; TYPE: AMINO ACID

```
;     TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-250-789A-5

  Query Match             38.7%;  Score 24;  DB 1;  Length 12;
  Best Local Similarity   66.7%;  Pred. No. 1.3e+02;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       4 LASSETGVG 12
           :: || |:|
Db       1 ISCSEAGLG 9


RESULT  11
US-08-250-789A-13
; Sequence 13, Application US/08250789A
; Patent No. 5635597
;  GENERAL INFORMATION:
;    APPLICANT:  Barrett, Ronald W.
;    APPLICANT:  Chernov-Rogan, Tania
;    APPLICANT:  Davis, Ann M.
;    TITLE OF INVENTION:  Peptides That Bind to Il-2 Receptors
;    NUMBER OF SEQUENCES:  194
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Townsend and Townsend Khourie and Crew
;      STREET:  One Market Plaza, Stewart Tower, Suite 2000
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94105
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/250,789A
;      FILING DATE:  27 MAY 1994
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  No. 5635597viel, Vernon A.
;      REGISTRATION NUMBER:  32,483
;      REFERENCE/DOCKET NUMBER:  16529A-52/1043
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-326-2400
;      TELEFAX:  415-326-2422
;  INFORMATION FOR SEQ ID NO:  13:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  12 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-250-789A-13


  Query Match             38.7%;  Score 24;  DB 1;  Length 12;
  Best Local Similarity   44.4%;  Pred. No. 1.3e+02;
  Matches    4;  Conservative    3;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       5 LASSETGVG 13
           :: || |:|
Db       1 ISCSEAGLG 9


RESULT  12
US-08-250-789A-45
; Sequence 45, Application US/08250789A
; Patent No. 5635597
;  GENERAL INFORMATION:
;    APPLICANT:  Barrett, Ronald W.
;    APPLICANT:  Chernov-Rogan, Tania
;    APPLICANT:  Davis, Ann M.
;    TITLE OF INVENTION:  Peptides That Bind to Il-2 Receptors
;    NUMBER OF SEQUENCES:  194
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Townsend and Townsend Khourie and Crew
;      STREET:  One Market Plaza, Stewart Tower, Suite 2000
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94105
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/250,789A
;      FILING DATE:  27-MAY-1994
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  No. 5635597viel, Vernon A.
;      REGISTRATION NUMBER:  32,483
;      REFERENCE/DOCKET NUMBER:  16529A-52/1043
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-326-2400
;      TELEFAX:  415-326-2422
;  INFORMATION FOR SEQ ID NO:  45:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  12 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-250-789A-45


  Query Match             38.7%;  Score 24;  DB 1;  Length 12;
  Best Local Similarity   44.4%;  Pred. No. 1.3e+02;
  Matches    4;  Conservative    3;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       5 LASSETGVG 13
           :: || |:|
Db       1 ISCSEAGLG 9


RESULT  12
US-08-448-398-9
; Sequence 9, Application US/08448398
; Patent No. 6187571
;  GENERAL INFORMATION:
;    APPLICANT:  PIGNARD, Annie
;    APPLICANT:  GREZES-BESSET, Bruno
;    APPLICANT:  GRISON, Ren
;    APPLICANT:  SCHNEIDER, Michel
;    TITLE OF INVENTION:  Use of a DNA sequence coding for an
;    TITLE OF INVENTION:  oxalic acid degrading protein as a selection gene
;    NUMBER OF SEQUENCES:  20
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  DENNISON, MESERVEY, PILLAR & CHRISNEY
;      STREET:  1745 Jefferson Davis Highway
;      CITY:  Arlington
;      STATE:  Virginia
;      COUNTRY:  USA
;      ZIP:  22202
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/448,398
;      FILING DATE:  October 24, 1995
;      CLASSIFICATION:  435
```

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US97/21219
; FILING DATE: 07-DEC-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/US97/21219
; FILING DATE: 07-DEC-1997
; AGENT/AGENT INFORMATION:
; NAME: SCHEINER, Britton
; REGISTRATION NUMBER: 34,159
; REFERENCE/DOCKET NUMBER: 95001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 412-1295
; TELEFAX: (212) 412-1161
; TELEX:
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 12 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-438-508-9

Query Match 38.7%; Score 24; DB 4; Length 12;
Best Local Similarity 66.7%; Pred. No. 1.3e+02;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 4 ILASSEIV 11
: : : : : : :
Db 3 IVASSSIV 9

RESULT 13
US-08-438-508-3
; Sequence 3, Application US/08438508
; Patent No. 5840724
; GENERAL INFORMATION:
; APPLICANT: Lawrence, Daniel A.
; APPLICANT: Strickland, David
; APPLICANT: Sharma, Joseph D.
; APPLICANT: Fay, William P.
; APPLICANT: Olson, Steven T.
; APPLICANT: Francis Chmura, Anne Marie
; APPLICANT: Daniel T. Kittman
; APPLICANT: Joel Barrett
; TITLE OF INVENTION: Peptide Mediated Enhancement of
; TITLE OF INVENTION: Thrombolysis: Methods and Compositions
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: TX
; COUNTRY: USA
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/438,508
; FILING DATE: Concurrently Herewith
; CLASSIFICATION: 530
; AGENT/AGENT INFORMATION:
; NAME: Parker, David L.
; REGISTRATION NUMBER: 32,165
; REFERENCE/DOCKET NUMBER: 7211/7000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (512) 418-3000
; TELEFAX: (512) 474-7577
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-438-508-3

Query Match 38.7%; Score 24; DB 4; Length 8;
Best Local Similarity 71.4%; Pred. No. 1.7e+05;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 4 ILASSEIV 10
: : : : : : :
Db 1 IVASSSI 7

RESULT 14
US-09-187-859-3857
; Sequence 3857, Application US/09187859A
; Patent No. 6358920
; GENERAL INFORMATION:
; APPLICANT: Blaschuk, Orest W.
; APPLICANT: Gour, Barbara J.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING NONCLASSICAL
; TITLE OF INVENTION: CADHERIN-MEDIATED FUNCTIONS
; FILE REFERENCE: 100086.407C1
; CURRENT APPLICATION NUMBER: US/09/187,859A
; CURRENT FILING DATE: 1998-11-06
; NUMBER OF SEQ ID NOS: 4052
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3857
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Representative cyclic modulating agent based on
; OTHER INFORMATION: cadherin-related neuronal receptor cell adhesion
; OTHER INFORMATION: recognition sequence
US-09-187-859-3857

Query Match 37.1%; Score 23; DB 4; Length 9;
Best Local Similarity 44.4%; Pred. No. 1.7e+05;
Matches 4; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KSEILASSE 9
: : : : : : :
Db 1 KKFTIDSSD 9

RESULT 15
US-09-140-149-13
; Sequence 13, Application Us/09140149
; Patent No. 6117680
; GENERAL INFORMATION:
; APPLICANT: Natesan, Sridaran
; APPLICANT: Gilman, Michael Z
; TITLE OF INVENTION: No. 6117680ei Compositions and Methods for Regulation of
; TITLE OF INVENTION: Transcription
; FILE REFERENCE: 363C
; CURRENT APPLICATION NUMBER: US/09/140,149
; CURRENT FILING DATE: 1998-08-26
; EARLIER APPLICATION NUMBER: 08/918,401
; EARLIER FILING DATE: 1997-08-26
; EARLIER APPLICATION NUMBER: 08/920,610
; EARLIER FILING DATE: 1997-08-27
; EARLIER APPLICATION NUMBER: 60/126,009
; EARLIER FILING DATE: 1998-07-29
; EARLIER APPLICATION NUMBER: PCT/US97/15219
; EARLIER FILING DATE: 1997-08-27
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 13
; LENGTH: 11
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:synthetic
; OTHER INFORMATION: peptide
DS 09 14:14:13

Query Match 37.1%; Score 23; DB 8; Length 11;
Best Local Similarity 55.6%; Pred. No. 1.9e-02;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 5 LASSEIGVG 13
| : | | |
Db 5 LAGAGAVGG 11

Search completed: August 26, 2002, 06:13:44
Job time: 1+1 sec